

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2004年1月8日 (08.01.2004)

PCT

(10) 国際公開番号  
WO 2004/004334 A1

(51) 国際特許分類: H04N 7/08, H04J 3/00

(21) 国際出願番号: PCT/JP2003/007639

(22) 国際出願日: 2003年6月17日 (17.06.2003)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:  
特願2002-185758 2002年6月26日 (26.06.2002) JP  
特願2003-81133 2003年3月24日 (24.03.2003) JP

(71) 出願人(米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府門真市大字門真1006番地 Osaka (JP).

(72) 発明者: および

(75) 発明者/出願人(米国についてのみ): 遠間 正真 (TOMA,Tadamasa) [JP/JP]; 〒560-0085 大阪府豊中市上新田4丁目8番C-408号 Osaka (JP). 松井 義徳 (MATSUI,Yoshinori) [JP/JP]; 〒630-0212 奈良県生駒市辻町341番1-604号 Nara (JP). 能登屋 陽司 (NOTOYA,Youji) [JP/JP]; 〒572-0055 大阪府寝屋川市御幸東町3丁目14番423号 Osaka (JP).

(74) 代理人: 新居 広守 (NII,Hiromori); 〒532-0011 大阪府大阪市淀川区西中島3丁目11番26号 新大阪末広センタービル3F 新居国際特許事務所内 Osaka (JP).

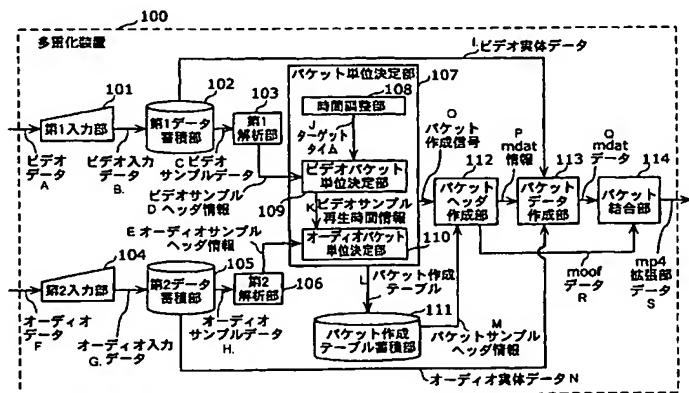
(81) 指定国(国内): CN, US.

(84) 指定国(広域): ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

[統葉有]

(54) Title: MULTIPLEXING DEVICE AND DEMULTIPLEXING DEVICE

(54) 発明の名称: 多重化装置および逆多重化装置



100...MULTIPLEXING DEVICE  
101...FIRST INPUT UNIT  
102...FIRST DATA ACCUMULATION UNIT  
103...FIRST ANALYSIS UNIT  
A...VIDEO DATA  
B...VIDEO INPUT DATA  
C...VIDEO SAMPLE DATA  
D...VIDEO SAMPLE HEADER INFORMATION  
E...AUDIO SAMPLE HEADER INFORMATION  
104...SECOND INPUT UNIT  
105...SECOND DATA ACCUMULATION UNIT  
106...SECOND ANALYSIS UNIT  
F...AUDIO DATA  
G...AUDIO INPUT DATA  
H...AUDIO SAMPLE DATA  
I...VIDEO SUBSTANTIAL DATA  
107...PACKET UNIT DECISION UNIT

108...TIME ADJUSTMENT UNIT  
J...TARGET TIME  
109...VIDEO PACKET UNIT DECISION UNIT  
K...VIDEO SAMPLE REPRODUCTION TIME INFORMATION  
110...AUDIO PACKET UNIT DECISION UNIT  
L...PACKET CREATION TABLE  
111...PACKET CREATION TABLE ACCUMULATION UNIT  
M...PACKET SAMPLE HEADER INFORMATION  
N...AUDIO SUBSTANTIAL DATA  
O...PACKET CREATION SIGNAL  
P...mdat INFORMATION  
Q...mdat DATA  
112...PACKET HEADER CREATION UNIT  
113...PACKET DATA CREATION UNIT  
114...PACKET CONCATENATION UNIT  
R...moof DATA  
S...mp4 EXTENDED PORTION DATA

(57) Abstract: A multiplexing device (100) includes a first input unit (101) for acquiring video data, a second input unit (104) for acquiring audio data, a first analysis unit (103) for analyzing video data to obtain video sample header information, a second analysis unit (106) for analyzing

[統葉有]

WO 2004/004334 A1



添付公開書類:  
— 國際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

---

audio data to obtain audio sample header information, a packet unit decision unit (107) for deciding a video data packetizing unit according to the video sample header information and after this, deciding an audio data packetizing unit aligned with the reproduction start time of the video sample arranged at the head of the video data packet unit, a packet header creation unit (112) for creating a packet header portion on the decided packet unit, a packet data creation unit (113) for creating a packet data portion on the decided packet unit, and a packet concatenation unit (114) for concatenating the created packet header portion and the packet data portion, thereby creating a packet.

(57) 要約: 多重化装置 100 は、ビデオデータを取得する第1入力部 101 と、オーディオデータを取得する第2入力部 104 と、ビデオデータを解析してビデオサンプルヘッダ情報を取得する第1解析部 103 と、オーディオデータを解析してオーディオサンプルヘッダ情報を取得する第2解析部 106 と、ビデオサンプルヘッダ情報に基づいてビデオデータのパケット化単位を決定した後に、ビデオデータのパケット化単位の先頭に配置されるビデオサンプルの再生開始時間と揃えてオーディオデータのパケット化単位を決定するパケット単位決定部 107 と、決定されたパケット単位でパケットヘッダ部を作成するパケットヘッダ作成部 112 と、決定されたパケット単位でパケットデータ部を作成するパケットデータ作成部 113 と、作成されたパケットヘッダ部とパケットデータ部とを結合してパケットを作成するパケット結合部 114 とを備える。

## 明細書

## 多重化装置および逆多重化装置

## 5 技術分野

本発明は、動画像データや音声データ等のメディアデータを多重化する多重化装置、および動画像データや音声データ等のメディアデータが多重化されたピット列を読み込んで逆多重化する逆多重化装置に関する。

## 背景技術

10 近年、通信ネットワークの大容量化および伝送技術の進歩により、インターネット上で、動画、音声、テキスト、あるいは、静止画等のマルチメディアコンテンツを含む動画像ファイルをパーソナルコンピュータに配信する動画配信サービスの普及が著しい。また、携帯端末等のいわゆる第3世代の移動体通信システムの規格の標準化を図ることを目的とする国際標準化団体 3GPP(Third Generation Partnership Project)で、無線による動画配信に関する規格として TS26.234(Transparent end-to-end packet switched streaming service)が定められる等の動きも見られ、動画配信サービスは、携帯電話機やPDA等の移動体通信端末への提供の拡大も見込まれている。

20 動画配信サービスにおいて、動画像ファイルを配信する際には、まず、多重化装置において、動画、静止画、音声およびテキスト等のメディアデータを取り込んで、メディアデータの再生に必要なヘッダ情報とメディアデータの実体データとを多重化して動画像ファイルデータを作成することが必要となるが、この動画像ファイルデータの多重化ファイルフォーマットとして、MP4ファイルフォーマットが注目されている。

このMP4ファイルフォーマットは、国際標準化団体である

ISO/IEC(International Standardization Organization/International Engineering Consortium) JTC1/SC29/WG 11において標準化が進められている多重化ファイルフォーマットであり、上記3GPPのTS26.234でも採用されていることから、広く普及するものと予想されている。

5 ここで、MP4ファイルのデータ構造について説明する。

MP4ファイルは、ボックスと呼ばれるオブジェクト単位でヘッダ情報やメディアデータの実体データが格納されており、複数のボックスを階層的に配列することによって構成される。

図1は、従来のMP4ファイルを構成するボックスの構造を説明する  
10 ための図である。

ボックス901は、ボックス901のヘッダ情報が格納されるボックスヘッダ部902と、ボックス901に含まれるデータ（例えば、そのボックスの下の階層のボックスや情報を記述するためのフィールド等）が格納されるボックスデータ格納部903とから構成される。

15 このボックスヘッダ部902は、ボックスサイズ904、ボックスタイプ905、バージョン906、フラグ907のフィールドを有している。

ボックスサイズ904は、このフィールドに割り当てられたバイトサイズも含めてボックス901全体のサイズ情報が記述されるフィールド  
20 である。

ボックスタイプ905は、ボックス901の種別を識別するための識別子が記述されるフィールドである。この識別子は、通常4つのアルファベット文字列によって表される。なお、以下、本明細書中において、この識別子によって各ボックスを示す場合がある。

25 バージョン906は、ボックス901のバージョンを示すバージョン番号が記述されるフィールドであり、フラグ907は、ボックス901

毎に設定されるフラグ情報が記述されるフィールドである。このバージョン906とフラグ907は、全てのボックス901に必須のフィールドではないので、これらのフィールドを有しないボックス901も存在しうる。

5 このような構造のボックス901が複数連なって構成されるMP4ファイルは、ファイルの構成に不可欠な基本部と、必要に応じて使用される拡張部とに大別することができる。まず、MP4ファイルの基本部について説明する。

図2は、従来のMP4ファイルの基本部を説明するための図である。

10 MP4ファイル910の基本部911は、ファイルヘッダ部912とファイルデータ部913とから構成される。

15 ファイルヘッダ部912は、ファイル全体のヘッダ情報、例えば、動画像（ビデオ）データの圧縮符号化方式等の情報が格納される部分であり、ファイルタイプボックス914とムービーボックス915とから構成される。

20 ファイルタイプボックス914は、“*ftyp*”の識別子で識別されるボックスであり、MP4ファイルを識別するための情報が格納される。MP4ファイルにどのようなメディアデータを格納するかについて、また、どのような圧縮符号化方式を用いた動画像（ビデオ）データや音声（オーディオ）データ等を格納するかについては、標準化団体やサービス事業者が独自に規定することができるため、MP4ファイルがどの規定に従って作成されたものであるかを識別するための情報を、このファイルタイプボックス914に格納する。

25 ムービーボックス915は、“*moov*”の識別子で識別されるボックスであり、ファイルデータ部913に格納される実体データのヘッダ情報、例えば、表示時間長等の情報が格納される。

ファイルデータ部 913 は、“mdat”の識別子で識別されるムービーデータボックス 916 によって構成される。なお、このファイルデータ部 913 の代わりに、この MP4 ファイル 910 とは異なる外部のファイルを参照することもできる。このように、外部のファイルを参照する場合には、MP4 ファイル 910 の基本部 911 は、ファイルヘッダ部 912 のみから構成されることになる。本明細書では、この外部ファイルの参照をする場合ではなく、MP4 ファイル 910 内に実体データを含む場合について説明する。

ムービーデータボックス 916 は、サンプルと称される単位でメディアデータの実体データを格納するボックスである。このサンプルとは、MP4 ファイルにおける最小のアクセス単位であり、MPEG(Moving Picture Experts Group)-4 Visual の圧縮符号化方式によって符号化したビデオデータの VOP(Video Object Plane) やオーディオデータのフレームに相当するものである。

ここで、従来における MP4 ファイルの基本部の構造について階層を掘り下げて、ムービーボックス 915 の構造を説明することとする。

図 3 は、従来の MP4 ファイルにおけるムービーボックスの構造を説明するための図である。

図 3 (a) に示すように、ムービーボックス 915 は、先に説明したボックスヘッダ部 902 とボックスデータ格納部 903 とから構成されている。そして、ボックスヘッダ部 902 を構成するボックスサイズ 904 のフィールドには、ムービーボックス 915 のサイズ情報が記述され（図 3 (a) では、“xxxx”とする。）、ボックスタイプ 905 のフィールドには、ムービーボックス 915 の識別子 “mooov” が記述される。

また、ムービーボックス 915 のボックスデータ格納部 903 には、

MP4ファイル910の基本部911のヘッダ情報が格納されるムービー<sup>5</sup>ヘッダボックス917や、ビデオトラックやオーディオトラック等、  
トラック毎のヘッダ情報が格納されるトラックボックス918等が格納  
されている。なお、ここにいうトラックとは、MP4ファイル910に  
10含まれる各メディアのサンプルデータ全体を意味し、動画像や音声やテ  
キスト等のトラックは、それぞれビデオトラック、オーディオトラック  
やテキストトラック等と称される。また、MP4ファイル910内に同一<sup>10</sup>メディアのデータが複数存在する場合は、同一メディアに対して複数  
のトラックが存在することになる。具体的に説明すると、例えば、MP  
15 4ファイル910内に2種類の動画像データが含まれている場合、2つ  
のビデオトラックが存在することになる。

ムービー<sup>15</sup>ヘッダボックス917も、先に説明したボックスヘッダ部9  
02とボックスデータ格納部903とから構成されており、ボックスヘ  
ッダ部902を構成するボックスサイズ904のフィールドには、ムー<sup>20</sup>ビー<sup>15</sup>ヘッダボックス917のサイズ情報が記述され(図3(a)では、  
“×××”とする。)、ボックスタイプ905のフィールドには、ムービー<sup>20</sup>  
ヘッダボックス917の識別子“mvhd”が記述される。そして、  
ムービー<sup>20</sup>ヘッダボックス917のボックスデータ格納部903には、M  
P4ファイル910の基本部911に含まれるコンテンツの再生に要す  
る時間長に関する情報等が格納される。

また、トラックボックス918のボックスヘッダ部902を構成する  
ボックスサイズ904のフィールドには、トラックボックス918のサ<sup>25</sup>イズ情報が記述され(図3(a)では、“××”とする。)、ボックスタイ  
プ905のフィールドには、トラックボックス918の識別子“trak”  
が記述される。そして、トラックボックス918のボックスデータ  
格納部903には、トラックヘッダボックス919が格納されている。

5 ト ラ ッ ク ヘ ッ ダ ボ ッ ク ス 9 1 9 は、ト ラ ッ ク 每 の ヘ ッ ダ 情 報 を 記 述 す  
る た め の フ ィ ー ル ド を 有 す る ボ ッ ク ス で あ り、" t k h d " の 識 別 子 に よ  
つ て 識 別 さ れ る。こ の ト ラ ッ ク ヘ ッ ダ ボ ッ ク ス 9 1 9 の ボ ッ ク ス デ タ  
格 納 部 9 0 3 に は、ト ラ ッ ク の 種 類 を 識 別 す る た め の ト ラ ッ ク I D を 記  
述 す る フ ィ ー ル ド や、ト ラ ッ ク の 再 生 に 要 す る 時 間 長 に 関 す る 情 報 等 が  
記 述 さ れ る。

10 このように、ムービーボックス915には、ボックス901が階層的に配列されており、“track”で識別されるトラックボックス918にビデオやオーディオ等のトラック毎のヘッダ情報が格納されている。そして、このトラックボックス918に含まれる下位のボックスにおいて、トラックのサンプル単位のヘッダ情報が格納されている。

図3(a)に示すムービーボックス915の構造をツリー状に示すと、図3(b)のような図が得られる。

すなわち、ムービーボックス915の下位のボックス群としてムービーヘッダボックス917、トラックボックス918が配列され、トラックボックス918の下位のボックス群としてトラックヘッダボックス919が配列されており、ボックス901が階層的に配置されていることがわかる。

MP4ファイルフォーマットの標準化当初、MP4ファイルは、上記基本部911のみから構成されていた。しかし、メディアデータの情報量が多くなると、サイズが大きくなってしまうので、ストリーミング再生への適用が難しい等の種々の問題があり、ヘッダボックスとデータボックスとの組が複数連なる拡張部の使用を加える改良がなされている。

25 図4は、従来における拡張部を含むMP4ファイルの構造を示す図である。

図4に示すように、上記改良が加えられたMP4ファイル920は、基本部911と拡張部921とから構成される。この拡張部921を含むMP4ファイル920では、全てのメディアデータを拡張部921に格納することができるので、MP4ファイル基本部911のムービーデータボックス916を省略することとしてもよい。

拡張部921は、所定の単位で区切られたパケット922が複数連なって構成される。

このパケット922は、ムービーフラグメントボックス923とムービーデータボックス916とが一対となって構成され、ムービーフラグメントとも称される。

ムービーデータボックス916は、上記区切られた所定の単位でトラック毎のサンプルを格納し、ムービーフラグメントボックス923は、このムービーデータボックス916に対応してヘッダ情報を格納するボックスであり、“m o o f”という識別子によって識別される。このムービーフラグメントボックス923の構造について、さらに詳しく説明する。

図5は、従来におけるムービーフラグメントボックスの構造を説明するための図である。

図5に示すように、ムービーフラグメントボックス923のボックスデータ格納部903には、ムービーフラグメントヘッダボックス924と複数のトラックフラグメントボックス925が格納されている。

ムービーフラグメントヘッダボックス924は、“m f h d”の識別子で識別されるボックスであり、ムービーフラグメントボックス923全体のヘッダ情報が格納される。

トラックフラグメントボックス925は、“t r a f”の識別子で識別されるボックスであり、トラック毎のヘッダ情報が格納される。

なお、通常 1 つのトラックのヘッダ情報に対して、1 つのトラックフラグメントポックス 925 が用意されるが、1 つのトラックのヘッダ情報に対して、複数のトラックフラグメントポックス 925 が用意されるとしてもよい。このように、1 つのトラックのヘッダ情報を複数のトラックフラグメントポックス 925 に分割して格納する際には、トラックフラグメントポックス 925 の先頭サンプルの復号時間が昇順となるように配列される。

そして、このトラックフラグメントポックス 925 のポックスデータ格納部 903 には、トラックフラグメントヘッダボックス 926 と 1 つ以上 のトラックフラグメントランボックス 927 が格納されている。

トラックフラグメントヘッダボックス 926 は、“t f h d” の識別子で識別されるボックスであり、トラックの種類を識別するためのトラック I.D を記述するフィールドや、サンプルの再生時間長等のデフォルト値に関する情報等を格納する。

トラックフラグメントランボックス 927 は、“t r u n” の識別子で識別されるボックスであり、サンプル単位のヘッダ情報を格納する。図 6 を用いて、このトラックフラグメントランボックス 927 について詳しく説明する。

図 6 は、従来におけるトラックフラグメントランボックス 927 の構造を説明するための図である。

フラグ 907 は、ポックス 901 毎に設定されるフラグ情報が記述されるフィールドであるが、ここでは、フラグ 907 に続いてデータオフセット 929 からサンプルコンポジションタイムオフセット 936 までの各フィールドがトラックフラグメントランボックス 927 に存在するか否かを示すフラグ情報が記述される。

サンプルカウント 928 は、トラックフラグメントランボックス 92

7 にどれだけの数のサンプルに関するヘッダ情報が格納されるかを示す  
情報が記述されるフィールドである。

データオフセット 929 は、トラックフラグメントランポックス 92  
7 にヘッダ情報が格納されているサンプルのうちトラックフラグメント  
5 ランポックス 927 の先頭に位置するサンプルの実体データが、組とな  
っているムービーデータボックス 916 のどこに格納されているかを示  
すポインタ情報が記述されるフィールドである。

先頭サンプルフラグ 930 は、トラックフラグメントランポックス 9  
27 の先頭サンプルがランダムアクセス可能なサンプルである場合に、  
10 後述するサンプルフラグ 935 のフィールドの値を上書きすることができる  
フィールドである。ここで、ランダムアクセスとは、例えば、MP  
4 ファイルの再生装置において、再生の途中でデータの再生位置を 10  
秒後に移動させたり、データの途中から再生を開始したりする処理動作  
を意味する。そして、ランダムアクセス可能なサンプルとは、ビデオサ  
15 ナップルのうち、MP4 ファイルの再生装置において、他のフレームのデ  
ータを参照することなく単独で復号化できるフレーム、すなわち画面内  
符号化フレーム（いわゆるイントラフレーム）を構成するサンプルを意味  
する。なお、オーディオサンプルでは、いずれのサンプルも単独で復  
号化することができるので、全てのオーディオサンプルがランダムア  
20 セス可能なサンプルといえる。

テーブル 931 は、サンプル毎のヘッダ情報を示すエントリ 932 が、  
サンプルカウント 928 において示される個数分集積されたものである。

エントリ 932 は、サンプル毎のヘッダ情報を示すフィールドの集まり  
25 であり、いずれのフィールドが含まれるかは、上記フラグ 907 によ  
って示される。エントリ 932 に含まれるフィールドには、サンプルのサ  
再生時間長が記述されるサンプルデュレーション 933、サンプルのサ

イズが記述されるサンプルサイズ 934、サンプルがランダムアクセス可能であるか否かを示すフラグ情報が記述されるサンプルフラグ 935、そして、双方向予測を用いたサンプルを扱うために、サンプルの復号時間と表示時間との差分値が記述されるサンプルコンポジションタイムオフセット 936 がある。

なお、これらのフィールドがエントリ 932 に含まれない場合は、各サンプルのヘッダ情報は、トラックフラグメントヘッダボックス 926 や、ムービーフラグメントボックス 915 内のムービーエクステンドボックス（識別子 “m v e x”）に、これらのフィールドのデフォルト値が記述されているので、これらのデフォルト値が使用される。

また、トラックフラグメントランボックス 927 には、復号時間の早いサンプルから順にヘッダ情報が記述される。従って、MP4 ファイルを再生する装置がサンプルのヘッダ情報を検索する際には、ファイル中の先頭のトラックフラグメントボックス 925 から順にトラックフラグメントヘッダボックス 926 内のトラック ID を参照することで、取得するトラックのヘッダ情報を含むトラックフラグメントボックス 925 を検索し、トラックフラグメントボックス 925 内においても、先頭のトラックフラグメントランボックス 927 から順にサンプルのヘッダ情報を検索することになる。

なお、この拡張部 921 を含む MP4 ファイル 920 の場合であっても、復号化時の初期化情報等、トラック全体に必要な情報は、ムービーボックス 915 に格納される。

続いて、このような構造を有する拡張部 921 を含む MP4 ファイルの構成例について説明する。

図 7 は、従来における拡張部を含む MP4 ファイルの拡張部の構成例を示す図である。

図7では、コンテンツの格納方法について2通りの例を示して説明することとし、コンテンツの再生時間長は、60秒であるとする。

図7(a)に示すMP4ファイル940は、基本部941および拡張部942の両方にメディアデータを格納する構成になっている。すなわち、基本部941のmdat\_1(符号945)に0~30秒までのメディアデータが格納され、拡張部942のmdat\_2(符号947)に30~45秒までのメディアデータが格納され、mdat\_3(符号949)に45~60秒までのメディアデータが格納されている。そして、mdat\_1(符号945)のヘッダ情報はmoo944に格納され、mdat\_2(符号947)のヘッダ情報はmoo\_f\_1(符号946)に格納され、mdat\_3(符号949)のヘッダ情報はmoo\_f\_2(符号948)に格納されている。

これに対して、図7(b)に示すMP4ファイル950は、拡張部952だけにメディアデータを格納する構成になっている。すなわち、基本部951は、fty953とmoo954とから構成されてmdatを含まず、拡張部952のmdat\_1(符号956)に0~30秒までのメディアデータが格納され、mdat\_2(符号958)に30~60秒までのメディアデータが格納されている。そして、mdat\_1(符号956)のヘッダ情報はmoo\_f\_1(符号955)に格納され、mdat\_2(符号958)のヘッダ情報はmoo\_f\_2(符号957)に格納されている。

ここで、上記MP4ファイルの拡張部がどのように作成されるかを図8~図10を用いて説明する。

図8は、従来の多重化装置の構成を示すブロック図である。

多重化装置960は、メディアデータを多重化してMP4ファイルの拡張部データを作成する装置である。ここでは、ビデオデータとオーデ

ィオデータとを多重化してMP4ファイルの拡張部データを作成するものとする。

第1入力部961はビデオデータを多重化装置960に取り込み、第1データ蓄積部962に蓄積させ、また、第2入力部964はオーディオデータを多重化装置960に取り込み、第2データ蓄積部965に蓄積させる。  
5

第1解析部963は、第1データ蓄積部962から1サンプルずつビデオデータを読み出して解析し、ビデオサンプルのヘッダ情報をパケット単位決定部967に出力する。また、第2階席部966は、第2データ蓄積部965から1サンプルずつオーディオデータを読み出して解析し、オーディオサンプルのヘッダ情報をパケット単位決定部967に出力する。このビデオサンプルヘッダ情報およびオーディオサンプルヘッダ情報には、サンプルのサイズや再生時間長を示す情報が含まれており、ビデオサンプルヘッダ情報には、ビデオサンプルがイントラフレームであるか否かを示す情報も含まれている。  
10  
15

パケット単位決定部967は、パケットに含まれるサンプル数が一定となるように、ビデオデータおよびオーディオデータのパケット単位を決定し、取得したサンプルヘッダ情報に基づいて各パケットのヘッダ情報を作成する。

20 図9に、従来におけるパケット単位決定部の処理動作フローを示す。ここで、1つのパケットに格納されるサンプルの数をNとし、この値は予め定められて、多重化装置960のメモリ等に保持されている。

まず、第1解析部963が1つのビデオサンプルを取得して(S901)、ビデオサンプルヘッダ情報をパケット単位決定部967に出力すると、パケット単位決定部967は、ビデオサンプルヘッダ情報をパケット作成テーブルに追加する(S902)。  
25

次に、パケット単位決定部 967 は、パケットに含まれるビデオサンプルの数を更新し (S903)、パケットに含まれるビデオサンプルの数が N になったかどうかを判定する (S904)。

ここで、パケットに含まれるビデオサンプルの数が N に満たない場合 5 (S904 の N○)、上記 S901～S903 までの処理が繰り返され、パケットに含まれるビデオサンプルの数が N になった場合 (S904 の Yes)、パケット単位決定部 967 は、N 個のビデオサンプルをパケット化して処理動作を終了する (S905)。

パケット単位決定部 967 は、同様に、オーディオについても上記 S 10 S901～S905 までの処理動作によって、オーディオサンプルのパケット化を行なう。

そして、全てのサンプルのパケット化が完了するまで、パケット単位決定部 967 は、このフローの処理動作を繰り返す。

図 10 に、従来におけるビデオサンプルのヘッダ情報を格納するパケット作成テーブルの一例を示す。このパケット作成テーブル 968a には、ビデオサンプル毎に、サンプルのサイズ、サンプルの再生時間長や、そのビデオサンプルがイントラフレームであるか否かを示す画面内符号化フレームフラグに関する情報が記述される。ここでは、パケットに格納される先頭のビデオサンプルは、サイズが 300 バイト、再生時間長が 30 ms、画面内符号化フレームでないことが示されており、2 番目のビデオサンプルは、画面内符号化フレームであることが示されている。そして、このパケット作成テーブル 968a は、パケット単位決定部 967 においてこれらの情報が順次追加され、1 パケットに含まれる最後のサンプルとなる N 番目まで作成されると、パケット作成テーブル蓄積部 25 968 に出力される。

再び図 8 を参照すると、続いて、パケット単位決定部 967 は、パケ

ット作成テーブル 968 a に N 個分のサンプルのヘッダ情報を記述した後、パケット作成テーブル 968 a をパケット作成テーブル蓄積部 968 に出力するとともに、パケットヘッダ作成部 969 にパケット作成信号を出力する。

5 パケットヘッダ作成部 969 は、パケット作成信号を取得すると、パケット作成テーブル蓄積部 968 に保持されているパケット作成テーブル 968 a からパケットサンプルヘッダ情報を読み出して m o o f データを作成する。また、パケットヘッダ作成部 969 は、作成した m o o f データをパケット結合部 971 に出力するとともに、パケットに含まれるサンプルの実体データが第 1 データ蓄積部 962 および第 2 データ蓄積部 965 のどこに格納されているかを示すポインタ情報と、サンプルのサイズ情報を含む m d a t 情報をパケットデータ作成部 970 に出力する。

10 パケットデータ作成部 970 は、取得した m d a t 情報に基づいて第 1 データ蓄積部 962 および第 2 データ蓄積部 965 からサンプルの実体データを読み出して m d a t データを作成し、 m d a t データをパケット結合部 971 に出力する。

15 そして、パケット結合部 971 は、 m o o f データと m d a t データとを結合させて、 1 パケット分の mp4 拡張部データを出力する。

20 最終的には、出力された 1 パケット分の mp4 拡張部データは、 MP4 ファイルを作成する装置に取り込まれ、順次作成される mp4 拡張部データが順番に並べられることによって、 MP4 ファイルの拡張部が作成される。その後、このファイル作成装置で、 MP4 ファイルの基本部と拡張部とが結合されることによって、 MP4 ファイルが作成されることになる。

25 しかしながら、このような従来の多重化装置によって多重化された M

MP4ファイルの拡張部を再生する際には、以下のような問題がある。

その1つとして、まず、従来の多重化装置では、パケットに含まれるサンプルの再生開始時間を考慮することなく多重化が行なわれるので、例えば、ある再生開始時間のビデオサンプルと同期が図られているオーディオサンプルが、ビデオサンプルと異なるパケットに格納される場合がある。そのため、MP4ファイルの再生装置側で、再生時のデータアクセスの効率が悪化するという問題がある。

また、従来の多重化装置では、パケットに含まれるサンプルの数を基準として多重化を行なうので、ランダムアクセス可能なサンプル、すなわちイントラフレームに相当するビデオサンプルをパケット内のどこに格納するかは、パケット毎にまちまちとなることが多い。そのため、MP4ファイルの再生装置側で、ランダムアクセス可能なサンプルを検索する際に、パケットに含まれる全てのビデオサンプルを検索しなければならず、サンプルの検索に要する計算量が膨大となってしまうという問題もある。

これらの問題について、図11を用いてさらに詳しく説明する。

図11は、従来における多重化装置の問題点を説明するための図である。

図11(a)では、再生時のデータアクセスの効率が悪化するという第1の問題を明らかにする。

各mdatに含まれるサンプルのヘッダ情報は、直前のmofに格納されており、mdat\_1に格納されている再生開始時間20sのビデオサンプルに関するヘッダ情報は、mof\_1に先頭サンプルとして格納されており、mdat\_10に格納されている再生開始時間20sのオーディオサンプルに関するヘッダ情報は、mof\_10に最終サンプルとして格納されている。

従って、MP4ファイルの再生装置が、コンテンツの再生時間20sの部分を再生しようとすれば、 $moo_f_1$ に格納されているビデオサンプルのヘッダ情報を取得してからオーディオサンプルのヘッダ情報を取得するまでに $moo_f_10$ まで検索しなければならず、データアクセスの効率が悪くなってしまう。

図11(b)では、ランダムアクセス可能なサンプルの検索に要する計算量が膨大となってしまうという第2の問題を明らかにする。

$mdat_1$ の最後に格納されているi番目のランダムアクセス可能なビデオサンプルに関するヘッダ情報は、 $moo_f_1$ に最終サンプルとして格納されており、 $mdat_3$ の最後に格納されているi+1番目のランダムアクセス可能なビデオサンプルに関するヘッダ情報は、 $moo_f_3$ に最終サンプルとして格納されている。

従って、MP4ファイルの再生装置が、ランダムアクセスを行なうとすれば、 $moo_f$ の最終サンプルまで検索しなければならず、検索に必要な計算量が膨大となってしまう。

さらに、これら第1および第2の問題に加えて、従来の多重化装置で作成されるMP4ファイルの拡張部の構成では、サンプルデータを取得するためのシークの回数が多くなるため、光ディスク再生機器等のシーク速度が遅い機器におけるランダムアクセス再生に適さないという問題もある。

この問題について、再び図11(b)を用いて説明する。 $moo_f_1$ のi番目のランダムアクセス可能なビデオサンプルにランダムアクセスしようとする場合、再生装置は、まず、i番目のランダムアクセス可能なビデオサンプルのヘッダ情報を取得するために、 $moo_f_1$ の先頭位置まで読み出しポインタを移動させ、 $moo_f_1$ 内を順に解析する。このとき、1回目のシークが必要となる。

その後、再生装置は、`m d a t _ 1`のどこに  $i$  番目のランダムアクセス可能なビデオサンプルの実体データが格納されているかを取得し、実体データの開始位置へ読み出しポインタを移動させる。このとき、 $i$  番目のランダムアクセス可能なビデオサンプルの実体データが`m d a t _ 5` 1 の終端に格納されているため、`m o o f _ 1` の先頭位置から連続的に読み出しポインタを移動させてサンプルの実体データを取得できず、2 回目のシークが必要となる。

すなわち、`m o o f _ 1` の先頭位置と実体データの開始位置に読み出しポインタを移動させる時にそれぞれシーク動作を行なうことになるので、再生装置がシーク速度の遅い機器である場合は、ランダムアクセス再生に時間がかかるてしまう。特に、この  $i$  番目のランダムアクセス可能なビデオサンプルと同期が図られているオーディオサンプル等の実体データが異なるパケット等、ビデオサンプルの実体データと離れて格納されている場合には、さらにシーク動作が必要となり、ランダムアクセス再生を迅速に行なうことが困難となる。

そこで、本発明は、これらの問題点に鑑みてなされたものであり、メディアデータの多重化ファイルが再生時のデータアクセスの効率に優れ、サンプルの検索に要する計算量が少なくなるようにメディアデータを多重化することができる多重化装置を提供することを目的とする。

また、多重化ファイルがシーク速度の遅い機器におけるランダムアクセス再生に適するようにメディアデータを多重化することができる多重化装置を提供することを目的とする。

さらに、このような多重化装置で多重化されたファイルを取得して、多重化分離することができる逆多重化装置を提供することを目的とする。

25

発明の開示

上記の目的を達成するために、本発明に係る多重化装置は、画像データと、音声データおよびテキストデータのうち少なくとも1つとを含むメディアデータをパケット多重化して多重化データを作成する多重化装置であって、前記メディアデータを取得するメディアデータ取得手段と、

5 前記メディアデータ取得手段が取得した前記メディアデータを解析して、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの最小のアクセス単位であるサンプルについて、サンプルの再生開始時間を示す再生開始時間情報を取得する解析手段と、前記解析手段が取得した前記再生開始時間情報に基づいて、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの各サンプルの再生開始時間を揃えて前記メディアデータをパケット化する単位を決定するパケット単位決定手段と、前記パケット単位決定手段が決定したパケット化単位で前記メディアデータのヘッダを格納するパケットヘッダ部を作成するパケットヘッダ部作成手段と、前記パケット単位決定手段が決定したパケット化単位で前記メディアデータの実体データを格納するパケットデータ部を作成するパケットデータ部作成手段と、前記パケットヘッダ部作成手段が作成したパケットヘッダ部と、前記パケットデータ部作成手段が作成したパケットデータ部とを結合してパケットを作成するパケット化手段とを備えることを特徴とする。

20 これによって、メディアデータに含まれる画像データと、音声データおよびテキストデータの再生開始時間が揃えられてパケットに格納されることとなるので、再生装置側で再生時におけるデータアクセスの効率を向上させることができる。

また、本発明に係る多重化装置は、前記画像データは、動画データであり、前記解析手段は、さらに、前記メディアデータ取得手段が取得した前記動画データを解析して、前記動画データが、画面内符号化サンプ

ルであることを示すイントラフレーム情報が含まれているサンプルを1つ以上含む場合に、前記イントラフレーム情報を取得し、前記パケット単位決定手段は、前記解析手段が前記イントラフレーム情報を取得した場合に、前記イントラフレーム情報と前記再生開始時間情報とに基づいて、前記メディアデータをパケット化する単位を決定し、前記イントラフレーム情報を含む前記動画データのサンプルを、前記パケット化単位の先頭に配置するのが好ましい。  
5

これによって、パケットに含まれる先頭のビデオサンプルは、イントラフレームのビデオサンプルとなるので、再生装置側でランダムアクセス時におけるサンプルの検索に要する計算量を大幅に削減することができる。  
10

さらに、本発明に係る多重化装置は、前記パケットデータ部作成手段は、前記パケット化単位に含まれる前記メディアデータのサンプルについて、サンプルの再生開始時間が昇順となるようにインタリーブして格納する前記パケットデータ部を作成するのがより好ましい。  
15

これによって、ビデオサンプルとオーディオサンプルとが再生開始時間が昇順となってm d a tに格納されるので、再生装置側でのランダムアクセス時におけるシーク動作の回数を少なくすることができ、シーク速度の遅い再生装置でも迅速なランダムアクセス再生を実現することができる。  
20

なお、本発明は、このような多重化装置として実現することができるだけでなく、このような多重化装置が備える特徴的な手段をステップとする多重化方法として実現したり、それらのステップをコンピュータに実行させるプログラムとして実現したりすることもできる。そして、そのようなプログラムは、CD-ROM等の記録媒体やインターネット等の伝送媒体を介して配信することができるのを言うまでもない。  
25

## 図面の簡単な説明

図1は、従来のMP4ファイルを構成するボックスの構造を説明するための図である。

図2は、従来のMP4ファイルの基本部を説明するための図である。

5 図3(a)は、従来のMP4ファイルにおけるムービーボックスの構造を説明するための図である。

図3(b)は、従来のMP4ファイルにおけるムービーボックスの構造をツリー状に示す図である。

10 図4は、従来における拡張部を含むMP4ファイルの構造を示す図である。

図5は、従来におけるムービーフラグメントボックスの構造を説明するための図である。

図6は、従来におけるトラックフラグメントボックスの構造を説明するための図である。

15 図7(a)は、従来における拡張部を含むMP4ファイルの第1の構成例を示す図である。

図7(b)は、従来における拡張部を含むMP4ファイルの第2の構成例を示す図である。

図8は、従来の多重化装置の構成を示すブロック図である。

20 図9は、従来におけるパケット単位決定部の処理動作を示すフロー図である。

図10は、従来におけるビデオサンプルのヘッダ情報を格納するパケット作成テーブルの一例を示す図である。

25 図11(a)は、従来における多重化装置の第1の問題点を説明するための図である。

図11(b)は、従来における多重化装置の第2の問題点を説明する

ための図である。

図12は、本発明の実施の形態1に係る多重化装置の機能的な構成を示すブロック図である。

図13は、多重化装置の処理動作を示すフロー図である。

5 図14は、ビデオパケット単位決定部の処理動作を示すフロー図である。

図15は、オーディオパケット単位決定部の処理動作を示すフロー図である。

10 図16(a)は、多重化装置が作成するMP4ファイル拡張部のデータ構造の第1例を示す図である。

図16(b)は、多重化装置が作成するMP4ファイル拡張部のデータ構造の第2例を示す図である。

図17は、本実施の形態2に係る多重化装置のパケット単位決定部の機能的な構成を示すブロック図である。

15 図18は、ビデオパケット単位決定部の第1の処理動作を示すフロー図である。

図19は、ビデオパケット単位決定部の第2の処理動作を示すフロー図である。

20 図20(a)は、多重化装置が作成するMP4ファイル拡張部のデータ構造の第1例を示す図である。

図20(b)は、多重化装置が作成するMP4ファイル拡張部のデータ構造の第2例を示す図である。

図21は、本実施の形態3に係る多重化装置のパケットデータ作成部の機能的な構成を示すブロック図である。

25 図22は、パケットデータ作成部の処理動作を示すフロー図である。

図23は、多重化装置が作成するMP4ファイル拡張部のデータ構造

の概略を示す図である。

図24は、多重化装置が作成するMP4ファイル拡張部のデータ構造の第1例を示す図である。

図25は、多重化装置が作成するMP4ファイル拡張部のデータ構造5の第2例を示す図である。

図26は、本実施の形態4に係る逆多重化装置の機能的な構成を示すブロック図である。

図27は、逆多重化装置の処理動作を示すフロー図である。

図28は、本発明に係る多重化装置の適用例を示す図である。

10

#### 発明を実施するための最良の形態

以下、本発明の実施の形態について、図面を参照しながら説明する。

なお、本実施の形態におけるビデオデータとして、MPEG-4 Visual の符号化データを用いることとし、本実施の形態におけるオーディオデータ15として、MPEG-4 Audio の符号化データを用いることとする。そして、本実施の形態では、主に、ビデオデータとオーディオデータとを多重化する装置について説明するが、テキストデータ等のその他のメディアデータの多重化について排除することを意図するものではない。

##### (実施の形態1)

20 まず、本発明の実施の形態1に係る多重化装置について、図12から図16を参照しながら説明する。

図12は、本発明の実施の形態1に係る多重化装置の機能的な構成を示すブロック図である。

この多重化装置100は、ビデオデータやオーディオデータを多重化25してMP4ファイルの拡張部データを作成する装置であり、第1入力部101、第1データ蓄積部102、第1解析部103、第2入力部10

4、第2データ蓄積部105、第2データ解析部106、パケット単位決定部107、パケット作成テーブル蓄積部111、パケットヘッダ作成部112、パケットデータ作成部113およびパケット結合部114を備える。

5 第1入力部101は、符号化されたビデオデータを画像符号化装置等から多重化装置100内に取り込むインターフェースであり、取得したビデオ入力データを順次、第1データ蓄積部102に蓄積させる。

第1データ蓄積部102は、ビデオ入力データを一時的に保持するキャッシュメモリやRAM(Random Access Memory)等である。

10 第1解析部103は、第1データ蓄積部102に保持されているビデオ入力データのうちビデオサンプル1つ分のデータであるビデオサンプルデータを読み出して解析し、ビデオサンプルのヘッダ情報を出力する処理部であり、CPUやメモリによって実現される。なお、この第1解析部103において出力されるビデオサンプルヘッダ情報には、ビデオサンプルのサイズ、再生時間長およびイントラフレームであるか否かを示す情報が含まれる。さらに、このビデオサンプルヘッダ情報には、双方向予測を用いたサンプルの場合、復号時間と表示時間の差分情報も含まれる。

20 第2入力部104は、符号化されたオーディオデータを音声符号化装置等から多重化装置100内に取り込むインターフェースであり、取得したオーディオ入力データを順次、第2データ蓄積部105に蓄積させる。

第2データ蓄積部105は、オーディオ入力データを一時的に保持するキャッシュメモリやRAM等である。

25 第2解析部106は、第2データ蓄積部105に保持されているオーディオ入力データのうちオーディオサンプル1つ分のデータであるオーディオサンプルヘッダ情報を出力する。

ディオサンプルデータを読み出して解析し、オーディオサンプルのヘッダ情報を出力する処理部であり、CPUやメモリによって実現される。なお、この第2解析部106において出力されるオーディオサンプルヘッダ情報には、オーディオサンプルのサイズおよび再生時間長を示す情報が含まれている。

5 パケット単位決定部107は、パケットに含まれるビデオサンプルおよびオーディオサンプルのヘッダ情報を集積させて、パケットに含まれるビデオサンプルの再生開始時間とオーディオサンプルの再生開始時間とが揃うように、ビデオデータおよびオーディオデータのパケット単位10を決定する処理部であり、CPUやメモリによって実現される。また、パケット単位決定部107は、決定したパケット単位分のサンプルヘッダ情報の集まりをパケット作成テーブルとしてパケット作成テーブル蓄積部111に出力するとともに、パケット単位の決定後にパケットヘッダの作成を指示するパケット作成信号をパケットヘッダ作成部112に出力する。そして、このパケット単位決定部107は、パケット単位を時間単位で調整する時間調整部108と、ビデオデータのパケット単位を決定するビデオパケット単位決定部109と、オーディオデータのパケット単位を決定するオーディオパケット単位決定部110とを備える。

時間調整部108は、パケットが定められた時間単位内に納まるよう20 に、パケットの終了時間を調整する処理部である。この時間調整部108は、まず、予め定められた時間（ターゲットタイム）をビデオパケット単位決定部109に出力する。なお、このターゲットタイムは、ユーザが指定することとしてもよい。この場合、多重化装置100は、キーボード等の入力装置を介してターゲットタイムの指定を取得し、入力装置25 から指定されたターゲットタイムを示すターゲットタイム入力信号が時間調整部108に出力されることとなる。

ビデオパケット単位決定部 109 は、第 1 解析部 103 からビデオサンプルヘッダ情報を取得してビデオデータのパケット単位を決定する処理部である。

このビデオパケット単位決定部 109 は、時間調整部 108 からターゲットタイムを、また、第 1 解析部 103 からビデオサンプルヘッダ情報を取得して、ビデオデータがターゲットタイム内のパケットに納まるように、各ビデオサンプルヘッダ情報に含まれる各ビデオサンプルの再生時間長をカウントしながら、パケットに含まれる最後のビデオサンプルのヘッダ情報をまで順次ビデオパケット作成テーブルに追加していく。

ビデオパケット単位決定部 109 は、パケットに含まれる最後のビデオサンプルのヘッダ情報をビデオパケット作成テーブルに追加すると、そのパケットに含まれる最初のビデオサンプルの再生開始時間とそのパケットに含まれるビデオサンプルの再生時間長の総和とを示すビデオサンプル再生時間情報をオーディオパケット単位決定部 110 に出力する。

オーディオパケット単位決定部 110 は、第 2 解析部 106 から取得したオーディオサンプルヘッダ情報を取得してオーディオデータのパケット単位を決定する処理部である。

このオーディオパケット単位決定部 110 は、ビデオパケット単位決定部 109 からビデオサンプル再生時間情報を、また、第 2 解析部 106 からオーディオサンプルヘッダ情報を取得して、パケットの先頭に、そのパケットに含まれる先頭のビデオサンプルの再生開始時間と同一または近似する再生開始時間のオーディオサンプルを配置し、各オーディオサンプルヘッダ情報に含まれる各オーディオサンプルの再生時間長をカウントしながら、そのパケットに含まれるオーディオサンプルの再生時間長の総和が、そのパケットに含まれるビデオサンプルの再生時間長の総和と同一または近似するように、そのパケットに含まれる最後のオ

ーディオサンプルを配置する。

なお、ここで、ビデオサンプルの再生開始時間と近似する再生開始時間のオーディオサンプルとは、ビデオサンプルの再生開始時間以降であって、最も早い再生開始時間のオーディオサンプル、または、ビデオサンプルの再生開始時間以前であって、最も遅い再生開始時間のオーディオサンプルを意味する。

その後、オーディオパケット単位決定部 110 は、パケットに含まれる先頭のオーディオサンプルから最後のオーディオサンプルまでのオーディオサンプルヘッダ情報を順次オーディオパケット作成テーブルに追加する。

パケット作成テーブル蓄積部 111 は、パケット単位決定部 107 から出力されるビデオパケット作成テーブルおよびオーディオパケット作成テーブルを一時的に保持するキャッシュメモリや RAM 等である。

パケットヘッダ作成部 112 は、パケットのヘッダ情報が格納されるパケットヘッダ部 (m o o f) を作成する処理部であり、CPU やメモリによって実現される。

このパケットヘッダ作成部 112 は、パケット単位決定部 107 からパケット作成信号を取得すると、パケット作成テーブル蓄積部 111 からパケット作成テーブルを参照してパケットサンプルヘッダ情報を読み出して m o o f データを作成し、パケット結合部 114 に出力する。

また、パケットヘッダ作成部 112 は、パケットに含まれるビデオサンプルおよびオーディオサンプルの実体データが、第 1 データ蓄積部 102 および第 2 データ蓄積部 105 のどこに格納されているかを示すポインタ情報や、サンプルのサイズを示すサンプルサイズ情報や、パケットデータ部 (m d a t) の作成を指示する信号が含まれる m d a t 情報をパケットデータ作成部 113 に出力する。

なお、このパケットヘッダ作成部 112 は、moof を作成する際に、  
例えば、AMR (Advanced Multi Rate CODEC) のような、データの途中で符  
号化レートの切替が発生する符号化方式によって符号化されたメディア  
データについて、符号化レートに応じてヘッダ情報を異なる traf に  
5 格納することもできる。

パケットデータ作成部 113 は、パケットの実体データが格納される  
パケットデータ部 (mdat) を作成する処理部であり、CPU やメモ  
リによって実現される。

このパケットデータ作成部 113 は、パケットヘッダ作成部 112 か  
10 ら mdat 情報を取得すると、mdat 情報に含まれるポインタ情報と  
サンプルサイズ情報とに基づいて、第 1 データ蓄積部 102 からパケッ  
トに含まれるビデオサンプルのビデオ実体データを読み出し、第 2 デー  
タ蓄積部 105 からパケットに含まれるオーディオサンプルのオーディ  
オ実体データを読み出して mdat データを作成し、パケット結合部 1  
15 14 に出力する。

パケット結合部 114 は、moof データと mdat データとを結合  
させて、1 パケット分の mp4 拡張部データを作成する処理部であり、  
CPU やメモリによって実現される。このパケット結合部 114 は、パ  
ケットヘッダ作成部 112 から moof データを取得し、パケットデータ  
20 作成部 113 から mdat データを取得して、moof データと mdat  
データとを結合させて 1 パケット分の mp4 拡張部データを作成し、  
順次作成した mp4 拡張部データを MP4 ファイルを作成する装置に出  
力する。

このように構成される多重化装置 100 において、MP4 ファイルの  
25 拡張部が作成される処理手順について図 13 を用いて説明する。

図 13 は、多重化装置 100 の処理動作を示すフロー図である。

まず、第1入力部101および第2入力部104は、多重化装置100内にそれぞれビデオデータおよびオーディオデータを取り込むと(S100)、第1入力部101はビデオ入力データを第1データ蓄積部102に蓄積させ、第2入力部104はオーディオ入力データを第2データ蓄積部105に蓄積させる。  
5

次に、第1解析部103は、第1データ蓄積部102からビデオサンプルデータを読み出して解析し、ビデオサンプルヘッダ情報をパケット単位決定部107のビデオパケット単位決定部109に出力する。そして、ビデオパケット単位決定部109は、第1解析部103から取得したビデオサンプルヘッダ情報を時間調整部108から取得したターゲットタイムとに基づいてビデオデータのパケット単位を決定する(S110)。なお、ビデオパケット単位決定部109がビデオデータのパケット単位を決定する処理動作については、詳しく後述する。  
10

その後、ビデオパケット単位決定部109は、パケット単位が決定されたパケットに含まれるビデオサンプルの再生時間情報をオーディオパケット単位決定部110に出力する(S120)。  
15

そして、オーディオパケット単位決定部110は、ビデオパケット単位決定部109から取得したビデオサンプルの再生時間情報に基づいて、オーディオデータのパケット単位を決定する(S130)。このとき、オーディオパケット単位決定部110は、パケットに含まれる先頭のオーディオサンプルの再生開始時間が、パケットに含まれる先頭のビデオサンプルの再生開始時間と同一またはこれに近似するように、パケット単位を決定する。  
20

オーディオパケット単位決定部110がオーディオデータのパケット単位を決定すると、パケット単位決定部107は、パケット作成テーブルをパケット作成テーブル蓄積部111に出力し、パケット作成信号を  
25

パケットヘッダ作成部 112 に出力する。

その後、パケットヘッダ作成部 112 は、決定された単位で m o o f データを作成してパケット結合部 114 に出力し、また、パケットデータ作成部 113 は、決定された単位で m d a t データを作成してパケット結合部 114 に出力し、パケット結合部 114 が m o o f データと m d a t データとを結合させて、決定された単位で 1 パケットを作成し (S 140)、1 パケット分の m p 4 拡張部データとして出力する。

1 パケットを作成し終えると、多重化装置 100 は、第 1 入力部 101 および第 2 入力部 104 から、まだ入力されるデータがあるか否かを判断する (S 150)。ここで、入力データがある場合 (S 150 の N o)、多重化装置 100 は、バッファメモリ、すなわち第 1 データ蓄積部 102、第 2 データ蓄積部 105 およびパケット作成テーブル蓄積部 111 に保持されているデータのうち、既にパケット化が終了したデータをクリアして (S 160)、上記 S 110 から S 150 までの処理動作を繰り返す。

一方、入力データがない場合 (S 150 の Y e s)、多重化装置 100 は、MP 4 ファイルの拡張部の作成処理を終了する。

このように、多重化装置 100 は、まずビデオデータのパケット単位を決定した後にオーディオデータのパケット単位を決定して、メディアデータの多重化を行なうことによって、MP 4 ファイルの拡張部を作成する。

ここで、図 13 のステップ S 110 において、ビデオパケット単位決定部 109 がビデオデータのパケット単位を決定する処理動作について詳しく説明する。

図 14 は、ビデオパケット単位決定部 109 の処理動作を示すフローリー図である。

このフローに先立ってビデオパケット単位決定部 109 は、時間調整部 108 からターゲットタイムを取得しておく。

そして、ビデオパケット単位決定部 109 は、第 1 解析部 103 からビデオサンプルヘッダ情報を取得すると (S111)、ビデオサンプルヘッダ情報をビデオパケット作成テーブルに追加する (S112)。

このとき、ビデオパケット単位決定部 109 は、ビデオサンプルヘッダ情報に含まれるビデオサンプルの再生時間長の合計、すなわちパケットに含まれるビデオデータの総再生時間が、先に取得したターゲットタイムになったか、あるいは、ターゲットタイムを超えたか否かを判定する (S113)。

パケットに含まれるビデオデータの総再生時間がターゲットタイムに至っていない場合 (S113 の No)、ビデオパケット単位決定部 109 は、次のビデオサンプルヘッダ情報を取得して (S111)、S112 と S113 の処理動作を繰り返す。

パケットに含まれるビデオデータの総再生時間がターゲットタイムに至っている場合 (S113 の Yes)、ビデオパケット単位決定部 109 は、ビデオパケット作成テーブルに最後に追加したビデオサンプルヘッダ情報が指示するビデオサンプルを、パケットに含まれる最後のビデオサンプルに決定し (S114)、パケット単位を決定する処理動作を終了する。

続いて、図 13 のステップ S130 において、オーディオパケット単位決定部 110 がオーディオデータのパケット単位を決定する処理動作について詳しく説明する。

図 15 は、オーディオパケット単位決定部 110 の処理動作を示すフロー図である。

このフローに先立って、オーディオパケット単位決定部 110 は、ビ

デオパケット単位決定部 109 からビデオサンプル再生時間情報を取得しておく。

そして、オーディオパケット単位決定部 110 は、第 2 解析部 106 からオーディオサンプルヘッダ情報を取得すると (S131)、先に取得 5 したビデオサンプル再生時間情報を参照して (S132)、パケットに含まれる先頭のビデオサンプルの再生開始時間を読み出し、パケットに含まれる先頭のビデオサンプルの再生開始時間と同一または近似する再生開始時間のオーディオサンプルを、そのパケットのオーディオ先頭サンプルに決定する (S133)。

10 オーディオパケット単位決定部 110 は、パケットに含まれるオーディオ先頭サンプルを決定すると、オーディオサンプルヘッダ情報を順次取得して (S134)、オーディオサンプルヘッダ情報をオーディオパケット作成テーブルに追加していく (S135)。

その後、オーディオパケット単位決定部 110 は、ビデオサンプル再生時間情報を参照して、パケットに含まれるビデオサンプルの再生時間長の総和を読み出し (S136)、パケットに含まれるオーディオサンプルの再生時間長の総和が、パケットに含まれるビデオサンプルの再生時間長の総和と同一または近似する値となるように、そのパケットに含まれる最後のオーディオサンプルを決定し (S137)、パケット単位を決定する処理動作を終了する。

このような多重化装置 100 による処理動作を経て作成される MP4 ファイルの拡張部は、再生装置側におけるデータアクセスの効率に優れている。その理由について、図 16 に多重化装置 100 が作成する MP4 ファイル拡張部のデータ構造の例を示して説明する。

25 図 16 (a) に示す MP4 ファイル拡張部 200 は、複数のパケットから構成され、MP4 ファイルの基本部に結合されている。

MP4ファイル拡張部200を構成する各パケットは、パケットヘッダ部のmoo\_fと、パケットデータ部のmdatから構成されている。ここで、パケット\_1は、MP4ファイル拡張部200の1番目のパケットであることを意味し、パケット\_1に含まれるmoo\_fは、moo\_f\_1、パケット\_1に含まれるmdatは、mdat\_1と示す。また、図16(a)の各mdat中に示す“V”は、ビデオサンプルであることを指し示すものであり、図16(a)の各mdat中に示す“A”は、オーディオサンプルであることを指し示すものである(以下、他の図においても同様とする。)。

MP4ファイル拡張部200のmdat\_1には、再生開始時間が20秒のビデオサンプルがビデオ先頭サンプルとして格納されており、同じく再生開始時間が20秒のオーディオサンプルがオーディオ先頭サンプルとして格納されている。また、mdat\_2にも、再生開始時間が30秒のビデオサンプルがビデオ先頭サンプルとして格納されており、同じく再生開始時間が30秒のオーディオサンプルがオーディオ先頭サンプルとして格納されている。

このように、1つのパケットにビデオサンプルとオーディオサンプルとを、各々の再生開始時間を揃えて格納することによって、再生装置側で、MP4ファイル拡張部200を再生する時に、データアクセスに要する計算量を大幅に削減することができる。

また、各メディアデータの再生開始時間が揃えられてパケットに格納されているので、任意の数のパケットでデータを分割して、MP4ファイルデータのサイズを所望のサイズに調整することもできる。

ここで、多重化装置100が作成するMP4ファイル拡張部は、図16(b)に示すデータ構造としてもよい。

図16(b)は、多重化装置100が作成するMP4ファイル拡張部

のデータ構造の第2例を示す図である。

図16(b)に示すMP4ファイル拡張部210のmdat\_1には、再生開始時間が20秒のビデオサンプルがビデオ先頭サンプルとして格納されており、mdat\_2には、再生開始時間が20秒のオーディオサンプルがオーディオ先頭サンプルとして格納されている。また、mdat\_3には、再生開始時間が30秒のビデオサンプルがビデオ先頭サンプルとして格納されており、mdat\_4には、再生開始時間が30秒のオーディオサンプルがオーディオ先頭サンプルとして格納されている。

このように、1つのパケットにビデオまたはオーディオのいずれか一方のデータを格納して、ビデオデータを格納するパケットと、再生開始時間が揃えられたオーディオデータを格納するパケットを交互に配列することによっても、再生装置側で、MP4ファイル拡張部200を再生する時に、データアクセスに要する計算量を大幅に削減することができる。

以上説明したように、本実施の形態1に係る多重化装置100によれば、各メディアデータの再生開始時間を揃えて、各メディアデータをパケット化するので、再生装置側におけるデータアクセスの効率化を図ることができる。

## 20 (実施の形態2)

次に、本発明の実施の形態2に係る多重化装置について、図17から図20を参照しながら説明する。

本実施の形態2に係る多重化装置は、主な構成要素において、上記実施の形態1に係る多重化装置100と共通するが、パケット単位決定部において特徴的な構成を備えており、この点において上記実施の形態1に係る多重化装置100と異なる。以下、この異なる点を中心に説明す

る。なお、上記実施の形態1と同一の構成要素については、同一の符号を用いることとし、説明を省略する。

図17は、本実施の形態2に係る多重化装置のパケット単位決定部の機能的な構成を示すブロック図である。

5 このパケット単位決定部117は、パケットに含まれるビデオサンプルおよびオーディオサンプルのヘッダ情報を集積させて、各々の再生開始時間が揃うように、かつ、パケットに含まれる先頭のビデオサンプルがイントラフレームとなるように、ビデオデータおよびオーディオデータのパケット単位を決定する処理部であり、時間調整部108と、ビデ  
10 オパケット単位決定部119と、オーディオパケット単位決定部110とを備える。

ビデオパケット単位決定部119は、第1解析部103からビデオサンプルヘッダ情報を取得してビデオデータのパケット単位を、時間またはイントラフレームのいずれかを基準に決定する処理部であり、時間基  
15 準単位調整部120と、Iフレーム基準単位調整部121とを備える。

時間基準単位調整部120は、時間調整部108から出力されるターゲットタイムに基づいてビデオデータのパケット単位を調整する処理部であり、各ビデオサンプルヘッダ情報の再生時間長をカウントして、パケットが定められた時間単位となるようにパケット単位を調整する。

20 Iフレーム基準単位調整部121は、第1解析部103から出力されるビデオサンプルヘッダ情報にイントラフレームであることを示す情報が含まれているか否かに基づいてビデオデータのパケット単位を調整する処理部であり、イントラフレームであることを示す情報が含まれているビデオサンプルヘッダ情報を取得すると、イントラフレームのビデオサンプルでパケット単位を切り替えて、次のパケットのビデオ先頭サンプルがイントラフレームのビデオサンプルとなるようにパケット単位を

調整する。

このように構成されるパケット単位決定部 117 を備えた本実施の形態 2 に係る多重化装置において、ビデオパケット単位決定部 119 がビデオデータのパケット単位を決定する処理動作について詳しく説明する。

5 図 18 は、ビデオパケット単位決定部 119 の処理動作を示すフロー図である。

このフローに先立って、ビデオパケット単位決定部 119 は、時間調整部 108 からターゲットタイムを取得して、時間基準単位調整部 120 に保持する。

10 そして、上記実施の形態 1 と同様に、ビデオパケット単位決定部 119 は、第 1 解析部 103 からビデオサンプルヘッダ情報を取得すると (S 201)、ビデオサンプルヘッダ情報をビデオパケット作成テーブルに追加する (S 202)。

15 このとき、ビデオパケット単位決定部 119 は、1 フレーム基準単位調整部 121 において、取得したビデオサンプルヘッダ情報にイントラフレームであることを示す情報が含まれているか否かを判定する (S 203)。

20 イントラフレームであることを示す情報が含まれている場合 (S 203 の Yes)、ビデオパケット単位決定部 119 は、時間基準単位調整部 120 において、パケットに含まれる全ビデオサンプルの総再生時間が、先に取得したターゲットタイムを超えていているか否かを判定する (S 205)。

25 ここで、イントラフレームであることを示す情報が含まれていない場合 (S 203 の No) またはターゲットタイムを超えていない場合 (S 205 の No)、ビデオパケット単位決定部 119 は、時間基準単位調整部 120 において、ビデオサンプルヘッダ情報に含まれるビデオサンプ

ルの再生時間長を加算することによって、パケットに含まれるビデオサンプルの再生時間長の総和を更新し(S204)、次のビデオサンプルヘッダ情報を取得して(S201)上記処理動作を繰り返す。

一方、ターゲットタイムを超えている場合(S205のYes)、ビデオパケット単位決定部119は、パケットに含まれる最後のビデオサンプルを、Iフレーム基準単位調整部121においてイントラフレームであると判定されたビデオサンプルの1つ前のビデオサンプルに決定し(S206)、ビデオデータのパケット単位決定の処理動作を終了する。

このようなビデオパケット単位決定部119の処理動作を経て作成されるMP4ファイルの拡張部は、パケットの先頭に格納されるビデオサンプルが必ずイントラフレームのビデオサンプルとなるので、再生装置側でランダムアクセス時にパケットの先頭のビデオサンプルから再生を開始することができるようになり、ランダムアクセス可能なビデオサンプルの検索に要する計算量を大幅に削減することができる。

また、パケットの先頭に格納されるビデオサンプルが必ずイントラフレームのビデオサンプルとなることによって、パケットヘッダ部(moo.f)では、ビデオトラックのヘッダ情報を格納するtrafの先頭に位置するtrunの先頭サンプルフラグフィールドにのみ、ランダムアクセス可能であることを示す情報を記述すればよく、各trunのサンプルフラグフィールドは、デフォルト値を使用することにより省略できるので、moo.fデータ作成時の負荷が軽減されるとともに、MP4ファイル全体のファイルサイズの削減を図ることもできる。

なお、この処理動作によると、ビデオデータに含まれるイントラフレーム同士の間隔が大きくなると、1パケットあたりの再生時間長が長くなる場合がある。そのため、パケット単位決定部117は、以下に述べるような処理動作としてもよい。

図19は、ビデオパケット単位決定部119の第2の処理動作を示すフロー図である。

上記第1の処理動作と同様に、このフローに先立って、ビデオパケット単位決定部119は、時間調整部108からターゲットタイムを取得して、時間基準単位調整部120に保持する。  
5

そして、ビデオパケット単位決定部119は、第1解析部103からビデオサンプルヘッダ情報を取得すると(S211)、ビデオサンプルヘッダ情報をビデオパケット作成テーブルに追加する(S212)。

このとき、ビデオパケット単位決定部119は、時間基準単位調整部120において、パケットに含まれる全ビデオサンプルの総再生時間が、  
10 先に取得したターゲットタイムを超えているか否かを判定する(S213)。

ターゲットタイムを超えている場合(S213のYes)、ビデオパケット単位決定部119は、パケットに含まれる最後のビデオサンプルを、  
15 今回取得したビデオサンプルヘッダ情報の1つ前のビデオサンプルヘッダ情報が指示するビデオサンプルに決定し(S214)、ビデオデータのパケット単位決定の処理動作を終了する。

一方、ターゲットタイムを超えていない場合(S213のNo)、ビデオパケット単位決定部119は、Iフレーム基準単位調整部121において、取得したビデオサンプルヘッダ情報にイントラフレームであることを示す情報が含まれているか否かを判定する(S215)。  
20

ここで、イントラフレームであることを示す情報が含まれている場合(S215のYes)、ビデオパケット単位決定部119は、パケットに含まれる最後のビデオサンプルを、Iフレーム基準単位調整部121においてイントラフレームであると判定されたビデオサンプルの1つ前のビデオサンプルに決定し(S214)、ビデオデータのパケット単位決定  
25

の処理動作を終了する。

他方、イントラフレームであることを示す情報が含まれていない場合 (S 215 の N○)、ビデオパケット単位決定部 119 は、時間基準単位調整部 120 において、ビデオサンプルヘッダ情報に含まれるビデオサンプルの再生時間長を加算することによって、パケットに含まれるビデオサンプルの再生時間長の総和を更新し (S 216)、次のビデオサンプルヘッダ情報を取得して (S 211) 上記処理動作を繰り返す。

このようなビデオパケット単位決定部 119 の第 2 の処理動作を経て作成される MP4 ファイルの拡張部は、所定の時間制限を設定してパケットを作成してパケットサイズを所望のサイズ以下に保ちつつ、イントラフレームのビデオサンプルが存在すれば、パケットの先頭に格納することができるので、再生装置側でランダムアクセス時にパケットの先頭のビデオサンプルについてのみランダムアクセス可能なビデオサンプルであるか否かを判定すればよくなり、ランダムアクセス可能なビデオサンプルの検索に要する計算量を削減することができる。

なお、ビデオパケット単位決定部 119 は、ビデオデータのパケット単位決定の処理動作を終了すると、ビデオサンプル再生時間情報をオーディオパケット単位決定部 110 に出力し、オーディオパケット単位 110 でオーディオデータのパケット単位決定の処理動作が行なわれるの 20 は、上記実施の形態 1 の場合と同様である。

このようなパケット単位決定部 117 による処理動作を経て作成される MP4 ファイルの拡張部は、再生装置側におけるランダムアクセス時の検索負荷を軽減させる。その理由について、図 20 に本実施の形態 2 に係る多重化装置が作成する MP4 ファイル拡張部のデータ構造の例を 25 示して説明する。

図 20(a)に示す MP4 ファイル拡張部 220 の m d a t \_ 1 には、

イントラフレームのビデオサンプルがビデオ先頭サンプルとして格納されており、`m d a t _ 2`にも同じくイントラフレームのビデオサンプルがビデオ先頭サンプルとして格納されている。

5 このように、イントラフレームのビデオサンプルを先頭のビデオサンプルとしてパケットに格納することによって、再生装置側でランダムアクセス時において、ランダムアクセス可能なビデオサンプルを取得するためにパケットの先頭のビデオサンプルのみを検索すれば足りるため、パケットに含まれる全てのビデオサンプルを検索する必要がなくなり、ランダムアクセス時のサンプル検索負荷を大幅に軽減することができる。

10 また、このとき、MP4ファイル拡張部220の`m o o f _ 1`および`m o o f _ 2`においても、ビデオトラックのヘッダ情報を格納する`t r a f`の先頭に位置する`t r u n`の先頭サンプルフラグフィールドにのみ、ランダムアクセス可能であることを示す情報を記述することによって、`m o o f _ 1`および`m o o f _ 2`のサイズを削減することもできる。

15 ここで、本実施の形態2に係る多重化装置が作成するMP4ファイル拡張部は、図20(b)に示すデータ構造としてもよい。

図20(b)に示すMP4ファイル拡張部230の`m d a t _ 1`には、イントラフレームのビデオサンプルがビデオ先頭サンプルとして格納されており、`m d a t _ 3`にも同じくイントラフレームのビデオサンプルがビデオ先頭サンプルとして格納されている。また、`m d a t _ 2`および`m d a t _ 4`には、オーディオサンプルが格納されている。

20 このように、1つのパケットにビデオまたはオーディオのいずれか一方のデータを格納して、ビデオデータを格納するパケットには、イントラフレームのビデオサンプルを先頭のビデオサンプルとして格納することによっても、再生装置側でランダムアクセス時におけるサンプル検索負荷を大幅に軽減することができる。

なお、これらMP4ファイル拡張部のデータ構造例のいずれにおいても、パケットに格納される先頭のビデオサンプルの再生開始時間と先頭のオーディオサンプルの再生開始時間とを揃えることによって、再生装置側でのデータアクセスに要する計算量を大幅に削減することができる。

5 以上説明したように、本実施の形態2に係る多重化装置によれば、ランダムアクセス可能なビデオサンプルを先頭のビデオサンプルとして、パケットを作成するので、再生装置におけるランダムアクセス時のサンプル検索に要する計算量を削減することができる。

(実施の形態3)

10 さらに、本発明の実施の形態3に係る多重化装置について、図21から図25を参照しながら説明する。

本実施の形態3に係る多重化装置は、主な構成要素において、上記実施の形態1および2に係る多重化装置と共通するが、パケットデータ作成部において特徴的な構成を備えており、この点において上記実施の形態1および2に係る多重化装置と異なる。以下、この異なる点を中心にして説明する。なお、上記実施の形態1および2と同一の構成要素については、同一の符号を用いることとし、説明を省略する。

図21は、本実施の形態3に係る多重化装置のパケットデータ作成部の機能的な構成を示すブロック図である。

20 このパケットデータ作成部130は、パケットデータ部(mdat)を、ビデオサンプルの実体データとオーディオサンプルの実体データとをインタリーブして格納することによって作成する処理部であり、mdat情報取得部131と、ビデオ実体データ読出部132と、オーディオ実体データ読出部133と、インタリーブ配列部134とを備える。

25mdat情報取得部131は、パケットヘッダ作成部112からmdat情報を取得して、パケットデータ作成部130を構成する他の各部

に実体データの読み出し指示や再生時間情報を出力する処理部である。

このmdat情報取得部131は、パケットヘッダ作成部112からmdat情報を取得するとmdat情報を解析して、ビデオサンプルおよびオーディオサンプルの再生開始時間と再生終了時間とを示す再生時間情報を取得し、この再生時間情報に基づいて、パケットに含まれる全てのビデオサンプルとオーディオサンプルとを再生開始時間が昇順となるように並び替える。

そして、mdat情報取得部131は、並び替えた順番に従って再生開始時間の若いサンプルから順に、ビデオ実体データ読み出し部132にビデオサンプルの実体データの読み出しを指示するビデオ読み出し指示を出力する、または、オーディオ実体データ読み出し部133にオーディオサンプルの実体データの読み出しを指示するオーディオ読み出し指示を出力する。このビデオ読み出しには、ビデオサンプルの実体データが第1データ蓄積部102のどこに格納されているかを示すポインタ情報とビデオサンプルのサイズ情報とが含まれており、オーディオ読み出しには、オーディオサンプルの実体データが第2データ蓄積部105のどこに格納されているかを示すポインタ情報とオーディオサンプルのサイズ情報とが含まれている。

ビデオ実体データ読み出し部132は、mdat情報取得部131からビデオ読み出し指示を取得して、第1データ蓄積部102からビデオ実体データを読み出す処理部である。このビデオ実体データ読み出し部132は、ビデオ読み出し指示に含まれるポインタ情報とサイズ情報とを参照して第1データ蓄積部102からビデオ実体データを読み出して、読み出したビデオ実体データをインターブ配列部134に出力する。

オーディオ実体データ読み出し部133は、mdat情報取得部131からオーディオ読み出し指示を取得して、第2データ蓄積部105からオーデ

ィオ実体データを読み出す処理部である。このオーディオ実体データ読み出部 133 は、オーディオ読み出指示に含まれるポインタ情報とサイズ情報とを参照して第 2 データ蓄積部 105 からオーディオ実体データを読み出して、読み出したオーディオ実体データをインタリーブ配列部 13

5 4 に出力する。

10 インタリーブ配列部 134 は、ビデオ実体データ読み出部 132 およびオーディオ実体データ読み出部 133 から出力される読み出ビデオデータおよび読み出オーディオデータを出力される順に逐次取得し、インタリーブして配列することによって mdat データを作成し、パケット結合部 1

10 14 に出力する処理部である。

15 このように構成されるパケットデータ作成部 130 を備えた本実施の形態 3 に係る多重化装置において、パケットデータ作成部 130 が mdat を作成する処理動作について詳しく説明する。

図 22 は、パケットデータ作成部 130 の処理動作を示すフロー図で

15 ある。

まず、パケットデータ作成部 130 は、mdat 情報取得部 131 において、パケットヘッダ作成部 112 から mdat 情報を取得する (S 301)。mdat 情報取得部 131 は、取得した mdat 情報を解析して、サンプルのポインタ情報とサイズ情報と再生時間情報を抽出する。

20 そして、mdat 情報取得部 131 は、抽出したサンプルの再生時間情報に基づいて、パケットに含まれる全てのビデオサンプルとオーディオサンプルとを再生開始時間が昇順となるように並び替える。続いて、mdat 情報取得部 131 は、並び替えた順番に従って再生開始時間の若いサンプルから順に、抽出したビデオサンプルのポインタ情報とサイズ情報とを含むビデオ読み出指示をビデオ実体データ読み出部 132 に出力する、または、抽出したオーディオサンプルのポインタ情報とサイズ情報

とを含むオーディオ読出指示をオーディオ実体データ読出部 133 に出力する。

ビデオ実体データ読出部 132 は、ビデオ読出指示を取得すると、ポインタ情報とサイズ情報とを参照して第 1 データ蓄積部 102 からビデオ実体データを読み出してインタリーブ配列部 134 に出力し、オーディオ実体データ読出部 133 は、オーディオ読出指示を取得すると、ポインタ情報とサイズ情報とを参照して第 2 データ蓄積部 105 からオーディオ実体データを読み出してインタリーブ配列部 134 に出力する (S302)。

10 インタリーブ配列部 134 は、読み出した実体データをビデオ実体データ読出部 132 およびオーディオ実体データ読出部 133 から受け取ると、受け取った順に逐次配列する (S303)。

ここで、インタリーブ配列部 134 は、ビデオ実体データとオーディオ実体データの全て、すなわち、1 パケットに格納される実体データの全ての配列が完了するまで、実体データの配列を続行する (S304 の N o.、S303)。

そして、1 パケットに格納される実体データの全ての配列が完了すると (S304 の Y e s)、インタリーブ配列部 134 は、配列した実体データを m d a t データとして、パケット結合部 114 に出力して (S305)、m d a t の作成の処理動作を終了する。

このようなパケットデータ作成部 130 の処理動作を経て作成される MP4 ファイルの拡張部は、シークに時間がかかる光ディスク機器等におけるランダムアクセス再生に適している。その理由について図 23 に本実施の形態 3 に係る多重化装置が作成する MP4 ファイル拡張部のデータ構造の概略を示して説明する。

図 23 に示す MP4 ファイル拡張部 240 は、4 ~ 8 秒までのコンテ

ンツデータを格納するパケット1、8～12秒までのコンテンツデータを格納するパケット2、12～16秒までのコンテンツデータを格納するパケット3というように、複数のパケットが配列されることで構成されている。

5 各パケットは、`moof 241`と`mdat 242`とから構成されており、`moof 241`には、ビデオトラックに関する`tfhd (V)`および`traf (V-1, V-2)`と、オーディオトラックに関する`tfhd (A)`および`traf (A-1, A-2)`とが格納されている。また、`traf (V-1)`と`traf (A-1)`に格納されるヘッダ情報が指し示すサンプルの実体データは、`mdat_1`に格納され、`traf (V-2)`と`traf (A-2)`に格納されるヘッダ情報が指し示すサンプルの実体データは、`mdat_2`に格納されている。そして、`mdat 242`には、ビデオサンプルの実体データとオーディオサンプルの実体データとが交互にインタリーブして格納されている。

10 15 このとき、再生装置側で、再生時間が4秒の位置から再生を開始するランダムアクセス処理に際して、`moof_1`の先頭位置に読み出しポインタを移動させれば、後は`moof_1`を解析して、読み出しポインタを連続的に移動させることにより`moof_1`に連続する`mdat_1`から再生に必要な実体データを取得することができる。

20 すなわち、このMP4ファイル拡張部240によれば、再生装置は、`moof_1`の先頭位置に読み出しポインタを移動させる1回のシーク動作だけで、ランダムアクセス再生を実現することができるので、シークに時間がかかる光ディスク機器等に有効といえる。

25 ここで、`mdat 242`において、ビデオサンプルの実体データの直後に格納されるオーディオサンプルの実体データは、直前のビデオサンプルの再生開始時間と揃えられているので、ビデオデータとオーディオ

データの同期再生は担保されている。図24に、MP4ファイル拡張部240のmdat\_1に実体データが格納されている様子を示す。

図24に示すように、mdat\_1の先頭に格納されているビデオサンプル1の再生開始時間は4000msであり、ビデオサンプル1の直後に格納されているオーディオサンプル1の再生開始時間は、4000msであり、ビデオサンプル1とオーディオサンプル1の再生開始時間は同一に揃えられている。

通常、ビデオサンプルとオーディオサンプルのサンプルレートは異なることが多いので、ここでは、ビデオサンプルの再生時間長は500msとし、オーディオサンプルの再生時間長は100msとする。

従って、MP4ファイル拡張部240のmdat\_1には、ビデオサンプル1の直後にオーディオサンプル1～5がインターブして格納され、その後に、ビデオサンプル2、オーディオサンプル6～10、ビデオサンプル3・・・の順に格納されることになる。

このとき、ビデオサンプル2の再生開始時間は、4500msであり、ビデオサンプル2の直後に格納されているオーディオサンプル6の再生開始時間も4500msであり、ビデオサンプルとそのビデオサンプル直後のオーディオサンプルの再生開始時間は、常に同一となるように揃えられている。

また、ビデオサンプルとオーディオサンプルのサンプルレートは異なるため、ビデオサンプルの再生開始時間とその直後のオーディオサンプルの再生開始時間とが同一とならない場合も生じうる。このような場合でも、ビデオサンプル直後のオーディオサンプルを、ビデオサンプルの再生開始時間と近似する再生開始時間有するオーディオサンプルとすることによって、ビデオデータとオーディオデータの同期再生を担保することができる。

図25は、MP4ファイル拡張部のm d a t \_ 1に実体データが格納されている様子を示す第2のデータ構造を示す図である。

図25に示すように、MP4ファイル拡張部250のm d a t \_ 1の先頭に格納されているビデオサンプル1の再生開始時間は、4000m  
5 sであり、ビデオサンプル1の直後に格納されているオーディオサンプル1の再生開始時間は、4050msであり、ビデオサンプル1の直後に格納されるオーディオサンプルとして、ビデオサンプル1の再生開始時間以降であって最も早い再生開始時間有するオーディオサンプル1  
10 が格納されている。

ここで、先に説明した場合と同様に、ビデオサンプルの再生時間長は500msとし、オーディオサンプルの再生時間長は100msとする。

従って、MP4ファイル拡張部250のm d a t \_ 1には、ビデオサンプル1の直後に、オーディオサンプル1～5がインタリーブして格納され、その後に、ビデオサンプル2、オーディオサンプル6～10、ビ  
15 デオサンプル3・・・の順に格納されることになる。

このとき、ビデオサンプル2の再生開始時間は、4500msであり、ビデオサンプル2の直後に格納されているオーディオサンプル6の再生開始時間は、4550msであり、ビデオサンプルとそのビデオサンプル直後のオーディオサンプルの再生開始時間は、常に近似するように揃  
20 えられている。

なお、ここで、ビデオサンプルの直後に格納されるオーディオサンプルとして、ビデオサンプルの再生開始時間以前であって最も遅い再生開始時間有するオーディオサンプルを格納することとしてもよい。この場合、ビデオサンプル1の直後に格納されるオーディオサンプル1は、  
25 3950msの再生時間を有することになる。

以上説明したように、本実施の形態3に係る多重化装置によれば、ビ

デオサンプルの直後に、ビデオサンプルの再生開始時間と同一または近似する再生開始時間有するオーディオサンプルを配置し、ビデオサンプルとオーディオサンプルとを再生開始時間が昇順となるようにインターブして *m d a t* に格納するので、シーク速度の遅い再生装置においても、迅速にランダムアクセス可能なデータ構造のMP4ファイル拡張部を作成することができる。

(実施の形態4)

続いて、本発明の実施の形態4に係る逆多重化装置について、図26および図27を参照しながら説明する。

図26は、本実施の形態4に係る逆多重化装置の機能的な構成を示すブロック図である。

逆多重化装置300は、上記実施の形態1、2および3に係る多重化装置で作成されたMP4ファイル拡張部を含むMP4ファイルデータを取得して解析し、メディアデータを逆多重化して再生データを出力する装置であり、ファイル入力部301、ファイルデータ蓄積部302、ヘッダ分離解析部303、*m o o v* 解析部304、*m o o f* 解析部305、*t r a f* 解析部306、*t r u n* 解析部307、RA検索部308およびサンプル取得部309を備えている。

ファイル入力部301は、MP4ファイルデータを取得するインターフェースであり、取得したMP4ファイルの入力データを順次、ファイルデータ蓄積部302に蓄積させる。

ファイルデータ蓄積部302は、MP4入力データを一時的に保持するキャッシュメモリやRAM等である。

ヘッダ分離解析部303は、ファイルデータ蓄積部302に保持されているMP4入力データのうちMP4ファイルのヘッダデータを読み出して解析し、MP4ファイルの基本部ヘッダの*m o o v*データと、拡張

部ヘッダのm o o f データとに分離して、それぞれm o o v 解析部304およびm o o f 解析部305に出力する処理部であり、CPUやメモリによって実現される。

5 m o o v 解析部304は、MP4ファイルのm o o v を解析して、メディアデータの符号化レートやコンテンツの再生時間長等、メディアデータの解析に必要なメディア情報を取得する処理部であり、CPUやメモリによって実現される。このm o o v 解析部は、取得したメディア情報をm o o f 解析部305に出力する。

10 m o o f 解析部305は、MP4ファイルのm o o f を、m o o v 解析部304から取得したメディア情報に基づいて解析し、トラック毎のヘッダデータであるt r a f データをt r a f 解析部306に出力する処理部であり、CPUやメモリによって実現される。

15 t r a f 解析部306は、MP4ファイルのt r a f を解析して、t r a f に含まれるサンプル毎のヘッダデータであるt r u n データをt r u n 解析部307に出力する処理部であり、CPUやメモリによって実現される。

20 t r u n 解析部307は、MP4ファイルのt r u n を解析して、t r u n 内の各フィールドに記述されている情報を取得して、サンプル取得部309にt r u n 解析情報を出力する処理部であり、CPUやメモリによって実現される。このt r u n 解析情報には、例えば、そのサンプルのサイズや、そのサンプルがファイルデータ蓄積部302のどこに格納されているかを示すデータオフセット情報や、さらにビデオサンプルの場合にはイントラフレームであることが否かを示すフラグ情報等が含まれている。

25 また、このt r u n 解析部307は、次に述べるRA検索部308から、ランダムアクセス後の再生開始位置を示し、再生の開始を指示する

出力信号である再生開始指示を取得すると、再生開始指示によって示される `trun` から順に解析して、サンプル取得部 309 に `trun` 解析情報を作成する。

RA 検索部 308 は、ランダムアクセス後の再生開始時間を示す目標再生時間情報を取得して、ビデオトラックに関するヘッダ情報を格納する先頭の `traf` 内の先頭の `trun` に含まれる先頭サンプルについての再生開始時間、およびイントラフレームであるかを示す情報である先頭サンプル情報を読み出して、ランダムアクセス後の再生開始位置となるビデオサンプルを検索する処理部であり、CPU やメモリによって実現される。この RA 検索部 308 は、ユーザからのランダムアクセス指示を受け付ける逆多重化装置 300 の入力装置から目標再生時間情報を取得すると、`trun` 解析部 307 から先頭サンプル情報を順次取得して、目標再生時間情報と同一または近似する再生開始時間を有するビデオサンプルを検索し、再生開始指示を `trun` 解析部 307 に出力する。

サンプル取得部 309 は、`trun` 解析情報に基づいて、サンプルの実体データを読み出して復号化し、再生データをディスプレイ等の表示装置に出力する処理部である。このサンプル取得部 309 は、`trun` 解析部 307 から `trun` 解析情報を取得すると、これに含まれるデータオフセット情報を参照して、ファイルデータ蓄積部 302 からサンプルの実体データを読み出す。ここで、`trun` 解析情報の取得開始をもって、再生開始が指示されたものとする。

このように構成される逆多重化装置 300 におけるランダムアクセス処理動作について図 27 を用いて説明する。

図 27 は、逆多重化装置 300 のランダムアクセス処理動作を示すフロー図である。なお、このフローに先立って、逆多重化装置 300 は、

入力装置を介してユーザからのランダムアクセス指示を受け付けているものとする。

まず、逆多重化装置300は、ファイル入力部301において、上記実施の形態1、2または3に係る多重化装置において作成されたMP4 5 ファイルのデータを取得すると(S400)、順次ファイルデータ蓄積部302に蓄積させていく。

次に、逆多重化装置300は、ヘッダ分離解析部303において、MP4ファイルのファイルヘッダ部のみを分離して解析し(S410)、さらに、基本部ヘッダと拡張部ヘッダとに分離して、moo解析部30 10 4において基本部ヘッダを解析し、mof解析部305において拡張部ヘッダを解析する(S420)。

続いて、逆多重化装置300は、mof解析部305において、拡張部ヘッダをさらに、トラック毎のヘッダに分離して、traf解析部306において、トラックフラグメント、すなわち、trafを解析する(S430)。このとき、逆多重化装置300は、traf解析部30 15 6において、トラックフラグメントをさらに分離して、trun解析部307において、trunを解析する。

ここで、逆多重化装置300は、RA検索部308において目標再生時間情報の入力があると、trun解析部307から先頭サンプル情報をRA検索部308に出力し、RA検索部308において、目標再生時間情報と同一または近似する再生開始時間が示されている先頭サンプル情報をあるか否かを判定する(S440)。

このとき、対象サンプルが見つからなければ(S450のNo)、逆多重化装置300は、RA検索部308において、ファイル内における格納順で次に配置された拡張部ヘッダにおける先頭サンプル情報を取得して、先に取得している目標再生時間情報と同一または近似する再生開始

時間が示されている先頭サンプル情報であるか否かを判定する (S 4 4 0)。

一方、対象サンプルが見つかれば (S 4 5 0 の Yes)、逆多重化装置 3 0 0 は、RA 検索部 3 0 8 において、再生開始指示を生成し、trun 5 解析部 3 0 7 に出力する。trun 解析部 3 0 7 は、RA 検索部 3 0 8 から再生開始指示を受けると、再生開始指示を受けた trun から順に、trun 解析情報をサンプル取得部 3 0 9 に出力する。ここで、再生開始指示を受けた trun とは、RA 検索部 3 0 8 において再生開始を指示されたサンプルを含む trun を指す。

10 その後、逆多重化装置 3 0 0 は、サンプル取得部 3 0 9 において、trun 10 解析情報に含まれるデータオフセット情報を参照して、ファイルデータ蓄積部 3 0 2 から対象サンプルの実体データを取得し (S 4 6 0)、復号化して再生データを出力してランダムアクセス処理動作を終了する。

以上説明したように、本実施の形態 4 に係る逆多重化装置 3 0 0 によれば、上記実施の形態 1、2 または 3 に係る多重化装置が作成する MP 4 ファイル拡張部を含む MP 4 ファイルについてランダムアクセス再生を行なう際に、各パケットの先頭に格納されているビデオサンプルのみを検索することによって、ランダムアクセス後の再生開始位置とすべきビデオサンプルを判定することができるので、ランダムアクセス時のサンプル検索負荷が大幅に軽減されることになる。

#### (適用例)

ここで、本発明に係る多重化装置の適用例について図 28 を用いて説明する。

図 28 は、本発明に係る多重化装置の適用例を示す図である。

25 本発明に係る多重化装置は、ビデオデータやオーディオデータ等のメディアデータを取得して多重化し、MP 4 ファイルデータを作成する録

画機能付き携帯電話機 403 やパーソナルコンピュータ 404 に適用されうる。また、本発明に係る逆多重化装置は、作成された MP4 ファイルデータを読み込んで再生する携帯電話機 407 に適用されうる。

ここで、録画機能付き携帯電話機 403 およびパーソナルコンピュータ 404 において作成された MP4 ファイルデータは、SD メモリカード 405 や DVD-RAM 406 等の記録媒体に格納されたり、通信ネットワーク 402 を介して画像配信サーバ 401 に送信されて、画像配信サーバ 401 から他の携帯電話機 407 等に配信されたりする。

このように、本発明に係る多重化装置および逆多重化装置は、画像配信システム等における MP4 ファイルの作成装置または再生装置として利用されるものである。

以上、本発明に係る多重化装置および逆多重化装置について、各実施の形態等に基づいて説明したが、本発明は、これらの実施の形態等に限定されるものではない。

例えば、上記各実施の形態では、ビデオデータとして、MPEG-4 Visual の符号化データを用いることとしたが、ビデオデータとして、MPEG-4 AVC (Advanced Video Coding) や H.263 等のその他の動画像圧縮符号化方式による符号化データを用いてもよい。なお、MPEG-4 AVC (Advanced Video Coding) や H.263 の符号化データでは、1 ピクチャが 1 サンプルに相当することになる。

同様に、オーディオデータとして、MPEG-4 Audio の符号化データを用いることとしたが、オーディオデータとして、G.726 等のその他の音声圧縮符号化方式による符号化データを用いてもよい。

また、上記各実施の形態では、ビデオデータとオーディオデータとを用いて説明しているが、テキストデータ等が含まれている場合でも、オーディオデータのパケット化と同じように処理することによって、本発

明の効果を得ることができる。

さらに、上記実施の形態2において、イントラフレーム毎にパケット化を行なうとする場合には、パケット単位決定部117の構成要素から時間基準単位調整部120を省略し、図18のステップS205の処理

5 を省略することとしてもよい。

またさらに、上記実施の形態3において、MP4ファイルの再生装置側で予め設定されているバッファモデルに従ってMP4ファイルが再生されることとなっている場合には、そのバッファモデルを満たすようにビデオサンプルのデータとオーディオサンプルのデータとをインタリーブしてmdatに格納することとしてもよい。ここで、バッファモデルとは、規格で定められた条件に従って符号化データが入力される場合に、その規格で定められたサイズのバッファを再生装置に持たせることで、バッファが空になる(アンダーフロー)、または、バッファから溢れる(オーバーフロー)ことなく、再生装置が復号化を行なうことができるこ

15 を保証するためのモデルである。

また、上記実施の形態1、2および3において、作成されるMP4ファイルの拡張部のmoofに格納するtrafの個数について言及していないが、moofに格納するtrafは、1つのトラックにつき1つのtrafを格納するのが好ましい。このようにすることで、トラック毎に、moof内の先頭trafのみを解析すれば、moofに格納されるトラックの全てのサンプルについてのヘッダ情報を取得することができるので、ヘッダ情報取得時の効率がさらに向上することとなる。

さらに、上記実施の形態1、2および3において、作成されるMP4ファイルの拡張部のmoofにヘッダ情報が格納されるサンプルの実体データは、moofに連続する1つのmdatに格納するとしているが、moofに連続する複数のmdatに分割して格納することとしてもよ

い。具体的に説明すると、`moo_f_1`にヘッダ情報が格納されるサンプルの実体データを、`mdat_1`、`mdat_2`、`mdat_3`の順に格納し、`moo_f_2`にヘッダ情報が格納されるサンプルの実体データを、`mdat_4`、`mdat_5`、`mdat_6`の順に格納するとし  
5 てもよい。

そして、上記実施の形態2および3では、パケット内に動画像データのイントラフレームが含まれる場合には、パケットの先頭に配置することとしているが、ランダムアクセスが可能であれば、P (Predictive) フレームやB (Bidirectionally predictive) フレーム等、イントラフレーム以外のビデオサンプルをパケットの先頭に配置することとしてもよい。  
10 以下、これについて、ビデオデータとして MPEG-4 AVC の符号化データを用いた場合を例に挙げて説明する。

MPEG-4 AVC では、イントラピクチャから復号化しても正しい復号結果を得られない場合がある。より詳しく説明すると、MPEG-4 AVC のイントラピクチャには、IDR (Instantaneous Decoder Refresh) ピクチャと、それ以外のピクチャ（以下、non-IDR イントラピクチャと称する。）の2種類があり、IDR ピクチャから復号化を開始すると、必ず正しい復号結果を得ることができるが、non-IDR イントラピクチャから復号化を開始すると、non-IDR イントラピクチャおよび表示順で non-IDR イントラピクチャ以降の複数枚のピクチャについて、正しい復号結果を得られないことがある。  
20

そのため、MPEG-4 AVC では、non-IDR イントラピクチャから正しい復号結果を得るためにには、どのピクチャから復号化を開始すればよいかを示す補助情報 (Recovery Point Supplemental Enhancement Information  
25 以下、“Recovery Point SEI” と称する。) を付加することができる。

例えば、`Pic_1`、`Pic_2`、`Pic_3`、`Pic_4`、`Pic`

5 で示される 5 枚のピクチャが、この順序でビデオデータに含まれ、

Pic\_5 が non-IDR イントラピクチャで、表示順で Pic\_5 および

Pic\_5 以降のピクチャを正しく復号化しようとすると、Pic\_1

から復号化を開始しなければならない場合、Pic\_1 の直前に、

5 Recovery Point SEI を配置することによって、ビデオデータ内における  
格納順で 4 枚後のピクチャである Pic\_5 、および、表示順でそれ以  
降のピクチャを正しく復号化するためには、Pic\_1 から復号化を開  
始する必要があることを示すことができる。

すなわち、この場合に、Pic\_1 は、ランダムアクセス可能なサン

10 プルであるといえるので、MPEG-4 AVC の符号化データの場合、IDR ピク  
チャまたは Recovery Point SEI が付加されたピクチャのサンプルを、ラ  
ンダムアクセス可能なサンプルとして、パケットの先頭に配置すること  
としてもよい。なお、Recovery Point SEI はイントラピクチャ以外のピ  
クチャに付加することもできる。

15 このとき、Recovery Point SEI が付加されたピクチャのサンプルと、  
Recovery Point SEI が付加されたピクチャから復号化を開始すること  
で初めて正しい復号結果を得られるようになるピクチャのサンプルとを同  
一パケットに格納することによって、サンプルデータ取得時の処理量を  
削減することができる。

20 さらに、IDR ピクチャと、Recovery Point SEI が付加されたピクチャ  
のサンプルとは、先頭サンプルフラグ 930 、あるいはサンプルフラグ  
935 における特定のフラグ値（以降、ノンシンクサンプルフラグと呼  
ぶ。）により識別することができる。MP4においては、ランダムアクセス  
可能なサンプルのうち、ランダムアクセスするサンプルと正しい復号  
結果が得られるサンプルとが一致するサンプルについてのみ、ノンシン  
クサンプルフラグを 0 にセットすることができる。このため、IDR ピク

チャのサンプルではノンシンクサンプルフラグを0とし、Recovery Point SEI が付加されたピクチャのサンプルではノンシンクサンプルフラグを1とすることにより、両者を識別することができる。

以上のような識別方法を用いることにより、IDR ピクチャと Recovery Point SEI が付加されたピクチャに限らず、互いに異なる性質をもつランダムアクセス可能なサンプルを識別することができる。実際には、以下のように使用することができる。

まず1つ目は、特定のサンプルのみを再生していくことにより、早送り再生を行う場合である。このときは、復号したサンプルをただちに表示できることが望ましいので、ノンシンクサンプルフラグが0であるサンプルのみを復号化し、再生することとする。

2つ目は、コンテンツの途中から再生を開始する、あるいは特定区間をスキップして次区間の再生を開始するような場合である。このとき、復号を開始するサンプルと正しい復号結果が得られるサンプルとが異なる可能性があるのは、再生開始時のみである。そこで、ノンシンクサンプルフラグが0であるサンプル、あるいはノンシンクサンプルフラグが1であるランダムアクセス可能なサンプルのどちらからでも再生を開始できることとする。

なお、このような格納方法は、MPEG-4 AVC の Recovery Point SEI の場合に限られず、復号化を開始するサンプルと、正しい復号結果が得られるサンプルとが異なる場合に適用することができ、例えば、MPEG2-Video における Open GOP (Group Of Pictures) のような構造に適用することができる。

さらに、サンプルがランダムアクセス可能であることを示す識別情報が付加されている際には、その識別情報によってランダムアクセス可能であることが示されているサンプルをパケットの先頭に配置することと

してもよい。

### 産業上の利用の可能性

本発明に係る多重化装置は、ビデオデータやオーディオデータ等のメ  
5 ディアデータを取得してMP4ファイルデータを作成し、記録媒体に格  
納するデジタルビデオカメラや録画機能付き携帯電話機等、または、作  
成したMP4ファイルデータをインターネットを介して配信するパソ  
ナルコンピュータやPDA等に用いるのに適し、本発明に係る逆多重化  
装置は、MP4ファイルデータをダウンロードして再生するパソナル  
10 コンピュータや携帯電話機等に用いるのに適している。

## 請求の範囲

1. 画像データと、音声データおよびテキストデータのうち少なくとも1つとを含むメディアデータをパケット多重化して多重化データを作成する多重化装置であって、

前記メディアデータを取得するメディアデータ取得手段と、

前記メディアデータ取得手段が取得した前記メディアデータを解析して、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの最小のアクセス単位であるサンプルについて、サンプルの再生開始時間を示す再生開始時間情報を取得する解析手段と、

前記解析手段が取得した前記再生開始時間情報に基づいて、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの各サンプルの再生開始時間を揃えて前記メディアデータをパケット化する単位を決定するパケット単位決定手段と、

15 前記パケット単位決定手段が決定したパケット化単位で前記メディアデータのヘッダを格納するパケットヘッダ部を作成するパケットヘッダ部作成手段と、

前記パケット単位決定手段が決定したパケット化単位で前記メディアデータの実体データを格納するパケットデータ部を作成するパケットデータ部作成手段と、

前記パケットヘッダ部作成手段が作成したパケットヘッダ部と、前記パケットデータ部作成手段が作成したパケットデータ部とを結合してパケットを作成するパケット化手段とを備える

ことを特徴とする多重化装置。

25

2. 前記パケット単位決定手段は、

前記パケット化単位の先頭に配置される前記画像データのサンプルの再生開始時間に、前記パケット化単位の先頭に配置される前記音声データおよび前記テキストデータのサンプルの再生開始時間を揃えることを特徴とする請求の範囲第1項記載の多重化装置。

5

3. 前記パケット単位決定手段は、

前記パケット化単位の先頭に配置される前記音声データおよび前記テキストデータのサンプルを、前記パケット化単位の先頭に配置される前記画像データのサンプルの再生開始時間以後であって、前記画像データのサンプルの再生開始時間に最も近い再生開始時間のサンプルとすることを特徴とする請求の範囲第2項記載の多重化装置。

10

4. 前記パケット単位決定手段は、

前記パケット化単位の先頭に配置される前記音声データおよび前記テキストデータのサンプルを、前記パケット化単位の先頭に配置される前記画像データのサンプルの再生開始時間以前であって、前記画像データのサンプルの再生開始時間に最も近い再生開始時間のサンプルとすることを特徴とする請求の範囲第2項記載の多重化装置。

20

5. 前記画像データは、動画データであり、

前記解析手段は、さらに、

前記メディアデータ取得手段が取得した前記動画データを解析して、前記動画データが、画面内符号化サンプルであることを示すイントラフレーム情報が含まれているサンプルを1つ以上含む場合に、前記イントラフレーム情報を取得し、

前記パケット単位決定手段は、

前記解析手段が前記イントラフレーム情報を取得した場合に、前記イントラフレーム情報と前記再生開始時間情報とに基づいて、前記メディアデータをパケット化する単位を決定する  
ことを特徴とする請求の範囲第1項記載の多重化装置。

5

6. 前記パケット単位決定手段は、  
前記イントラフレーム情報を含む前記動画データのサンプルを、前記パケット化単位の先頭に配置する  
ことを特徴とする請求の範囲第5項記載の多重化装置。

10

7. 前記パケット単位決定手段は、  
前記パケット化単位の先頭に配置される前記イントラフレーム情報を含む前記動画データのサンプルの再生開始時間に、前記パケット化単位の先頭に配置される前記音声データおよび前記テキストデータのサンプルの再生開始時間を揃える  
ことを特徴とする請求の範囲第6項記載の多重化装置。

8. 前記パケットデータ部作成手段は、  
前記パケット化単位に含まれる前記メディアデータのサンプルについて、サンプルの再生開始時間が昇順となるようにインタリーブして格納する前記パケットデータ部を作成する  
ことを特徴とする請求の範囲第1項記載の多重化装置。

9. 前記パケットデータ部作成手段は、  
前記パケット化単位に含まれる前記メディアデータのサンプルを、予め設定されている規定を満たすようにインタリーブして格納する前記パ

ケットデータ部を作成する

ことを特徴とする請求の範囲第8項記載の多重化装置。

10. 画像データと、音声データおよびテキストデータのうち少なく  
5 とも1つとを含むメディアデータをパケット多重化して多重化データを作成する多重化方法であって、

前記メディアデータを取得するメディアデータ取得ステップと、

前記メディアデータ取得ステップにおいて取得した前記メディアデータを解析して、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの最小のアクセス単位であるサンプルについて、サンプルの再生開始時間を示す再生開始時間情報を取得する解析ステップと、

前記解析ステップにおいて取得した前記再生開始時間情報に基づいて、前記メディアデータに含まれる前記画像データ、音声データおよびテキ  
15 ストデータの各サンプルの再生開始時間を揃えて前記メディアデータをパケット化する単位を決定するパケット単位決定ステップと、

前記パケット単位決定手ステップにおいて決定したパケット化単位で前記メディアデータのヘッダを格納するパケットヘッダ部を作成するパケ  
20 ットヘッダ部作成ステップと、

前記パケット単位決定ステップにおいて決定したパケット化単位で前記メディアデータの実体データを格納するパケットデータ部を作成するパケ  
25 ットデータ部作成ステップと、

前記パケットヘッダ部作成ステップにおいて作成したパケットヘッダ部と、前記パケットデータ部作成ステップにおいて作成したパケットデータ部とを結合してパケットを作成するパケット化ステップとを含むことを特徴とする多重化方法。

11. 前記パケット単位決定ステップにおいて、

前記パケット化単位の先頭に配置される前記画像データのサンプルの  
再生開始時間に、前記パケット化単位の先頭に配置される前記音声データ  
5 および前記テキストデータのサンプルの再生開始時間を揃える

ことを特徴とする請求の範囲第10項記載の多重化方法。

12. 前記画像データは、動画データであり、

前記解析ステップにおいて、さらに、

10 前記メディアデータ取得ステップにおいて取得した前記動画データを  
解析して、前記動画データが、画面内符号化サンプルであることを示す  
イントラフレーム情報が含まれているサンプルを1つ以上含む場合に、  
前記イントラフレーム情報を取得し、

前記パケット単位決定ステップにおいて、

15 前記解析ステップにおいて前記イントラフレーム情報を取得した場合  
に、前記イントラフレーム情報と前記再生開始時間情報とに基づいて、  
前記メディアデータをパケット化する単位を決定する  
ことを特徴とする請求の範囲第10項記載の多重化方法。

20 13. 前記パケット単位決定ステップにおいて、

前記イントラフレーム情報を含む前記動画データのサンプルを、前記  
パケット化単位の先頭に配置する

ことを特徴とする請求の範囲第12項記載の多重化方法。

25 14. 前記パケット単位決定ステップにおいて、

前記パケット化単位の先頭に配置される前記イントラフレーム情報を

含む前記動画データのサンプルの再生開始時間に、前記パケット化単位の先頭に配置される前記音声データおよび前記テキストデータのサンプルの再生開始時間を揃える

ことを特徴とする請求の範囲第13項記載の多重化方法。

5

15. 前記パケットデータ部作成ステップにおいて、

前記パケット化単位に含まれる前記メディアデータのサンプルについて、サンプルの再生開始時間が昇順となるようにインタリーブして格納する前記パケットデータ部を作成する

10 ことを特徴とする請求の範囲第10項記載の多重化方法。

16. 画像データと、音声データおよびテキストデータのうち少なくとも1つとを含むメディアデータをパケット多重化して多重化データを作成する多重化装置のためのプログラムであって、

15 前記メディアデータを取得するメディアデータ取得ステップと、

前記メディアデータ取得ステップにおいて取得した前記メディアデータを解析して、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの最小のアクセス単位であるサンプルについて、サンプルの再生開始時間を示す再生開始時間情報を取得する解析ステップと、

前記解析ステップにおいて取得した前記再生開始時間情報に基づいて、前記メディアデータに含まれる前記画像データ、音声データおよびテキストデータの各サンプルの再生開始時間を揃えて前記メディアデータをパケット化する単位を決定するパケット単位決定ステップと、

25 前記パケット単位決定手ステップにおいて決定したパケット化単位で前記メディアデータのヘッダを格納するパケットヘッダ部を作成するパ

ケットヘッダ部作成ステップと、

前記パケット単位決定ステップにおいて決定したパケット化単位で前記メディアデータの実体データを格納するパケットデータ部を作成するパケットデータ部作成ステップと、

5 前記パケットヘッダ部作成ステップにおいて作成したパケットヘッダ部と、前記パケットデータ部作成ステップにおいて作成したパケットデータ部とを結合してパケットを作成するパケット化ステップとを含む多重化方法における各ステップをコンピュータに実行させる  
ことを特徴とするプログラム。

10

17. 画像データと、音声データおよびテキストデータのうち少なくとも1つとを含むメディアデータが所定のパケットの単位で多重化されている多重化データを取得して逆多重化する逆多重化装置であって、

前記多重化データを取得する多重化データ取得手段と、

15 前記多重化データ取得手段が取得した前記多重化データを解析して、前記パケットのヘッダ部を前記多重化データから分離して取得する解析分離手段と、

前記多重化データの逆多重化の開始位置を変更する、あるいは前記多重化データの途中から逆多重化を開始する処理であるランダムアクセスを実行する際に、前記解析分離手段が分離したパケットヘッダ部の先頭に配置されている前記画像データのサンプルのヘッダのみを検索して、前記パケットに含まれる前記画像データのサンプルが画面内符号化サンプルであることを示すイントラフレーム情報が含まれているか否かを判定するランダムアクセス検索手段とを備える  
ことを特徴とする逆多重化装置。

図1

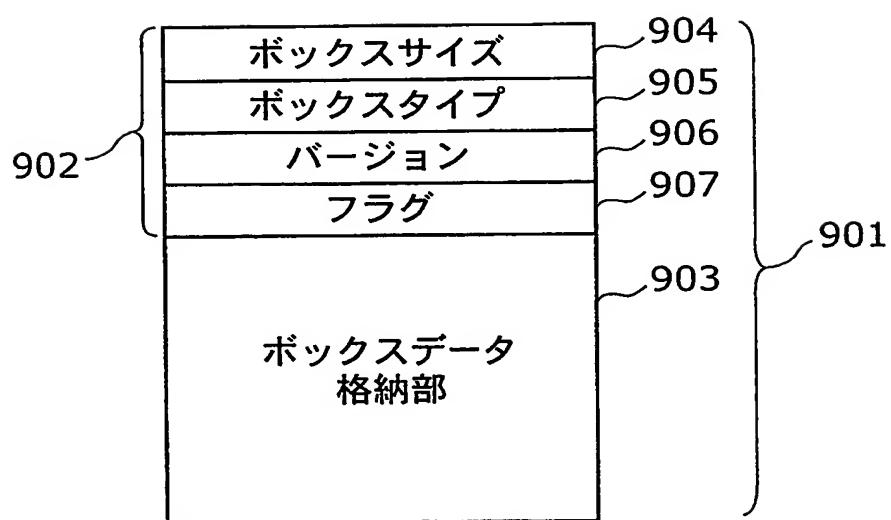


図2

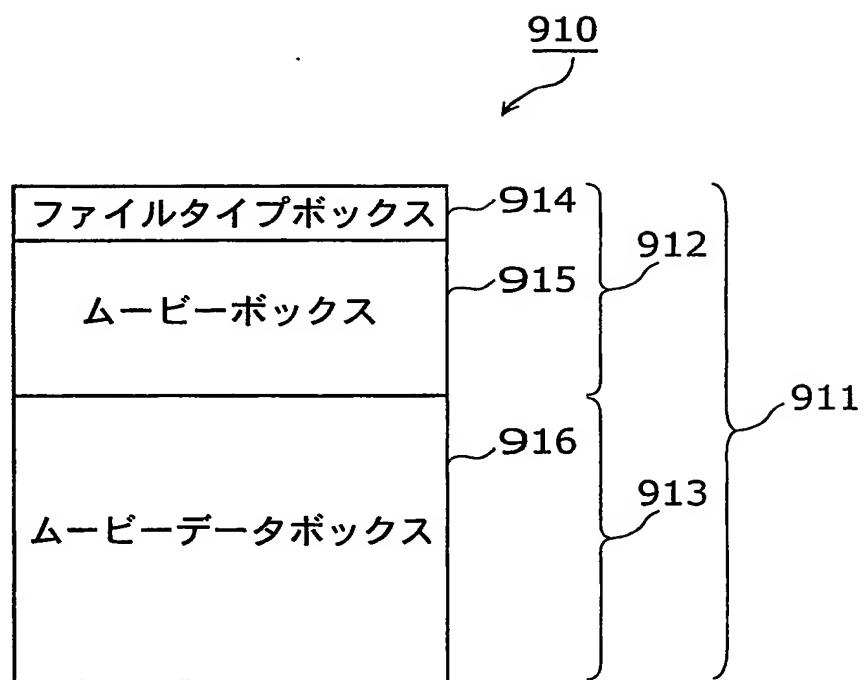


図3

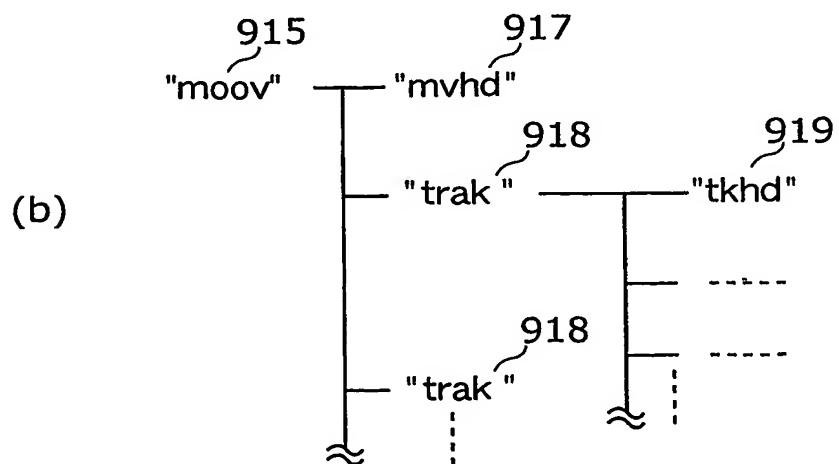
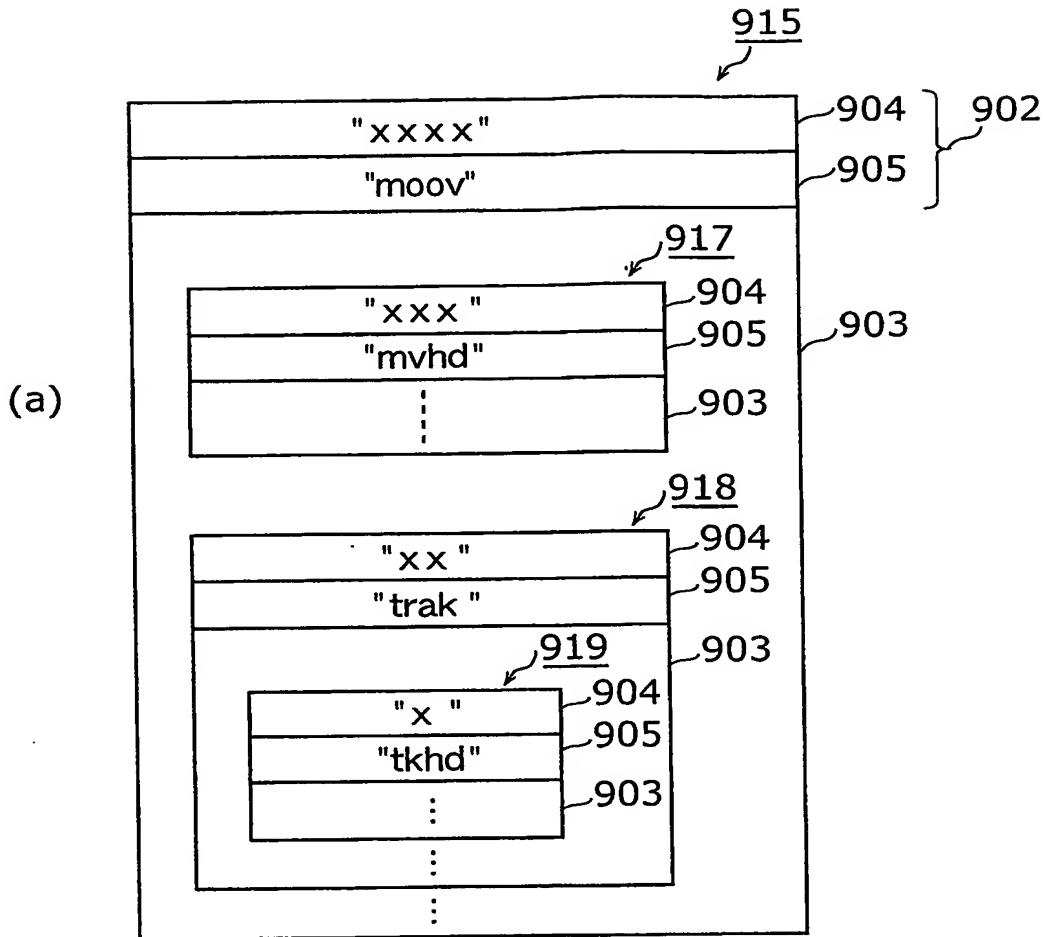


図4

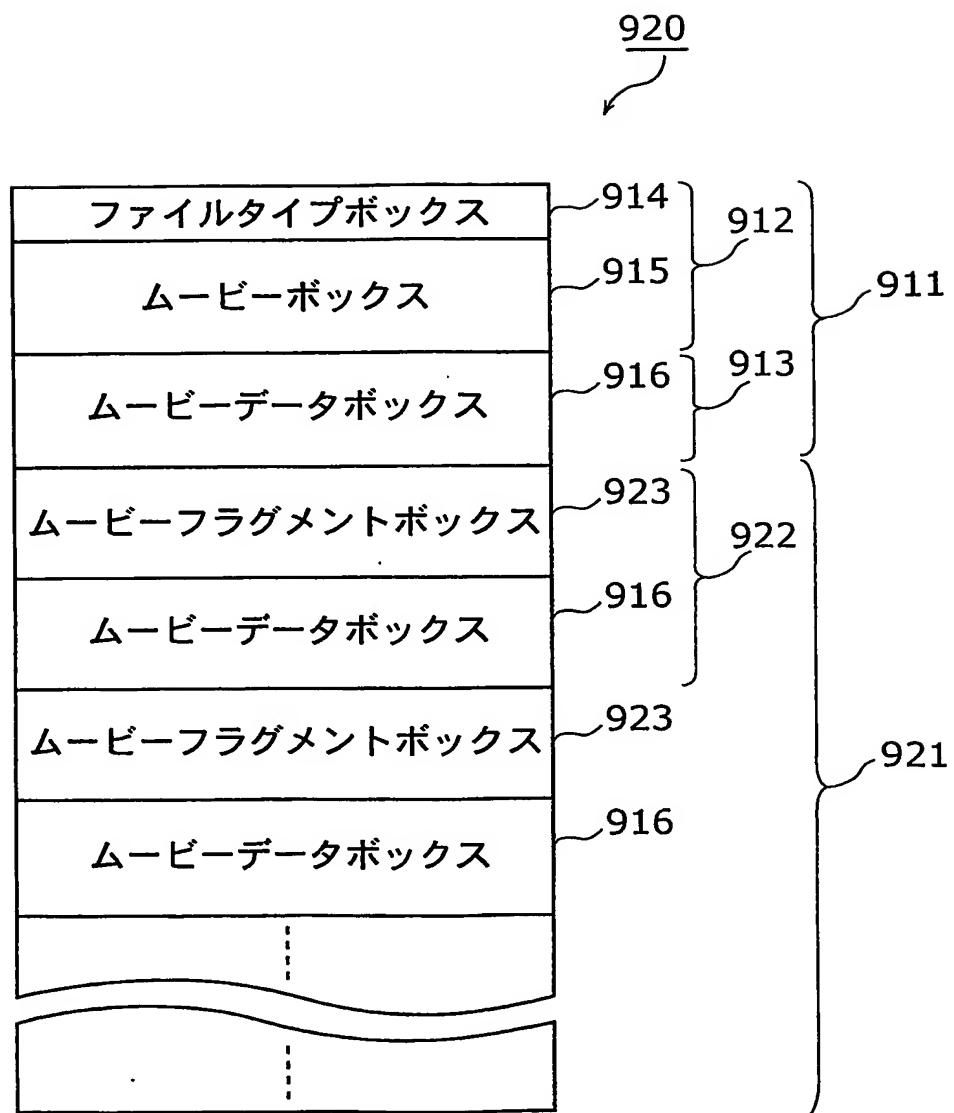


図5

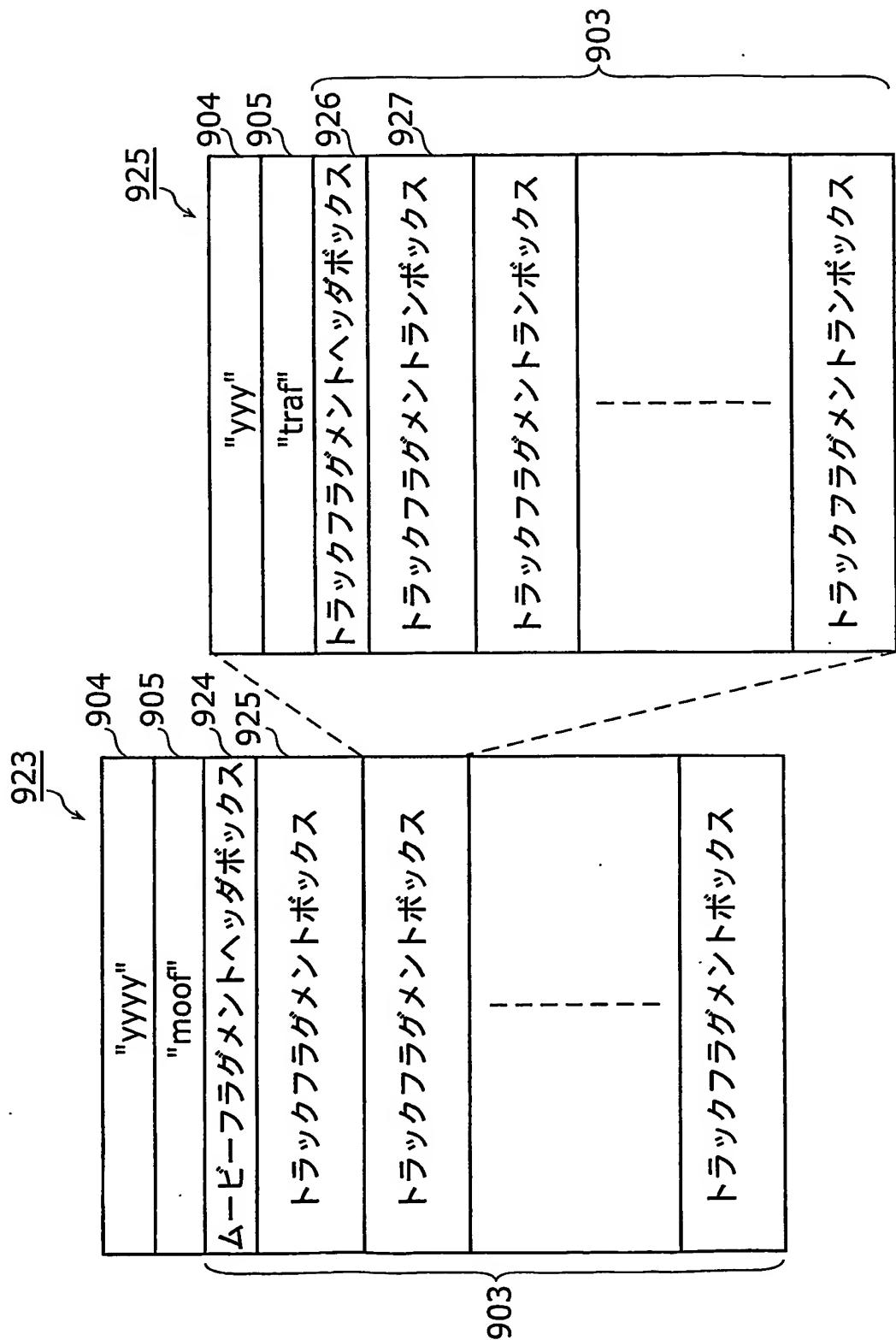


図6

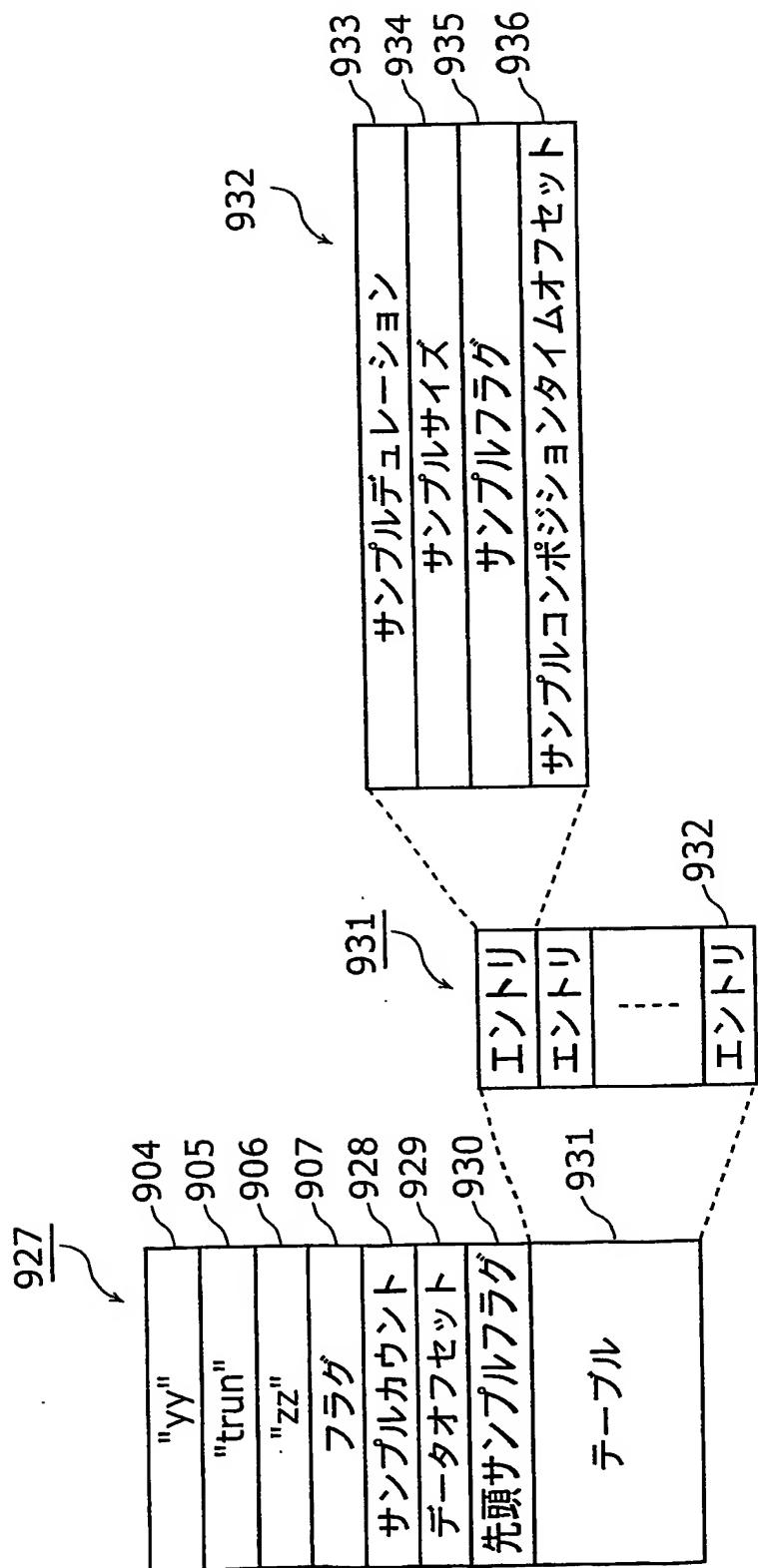


図7

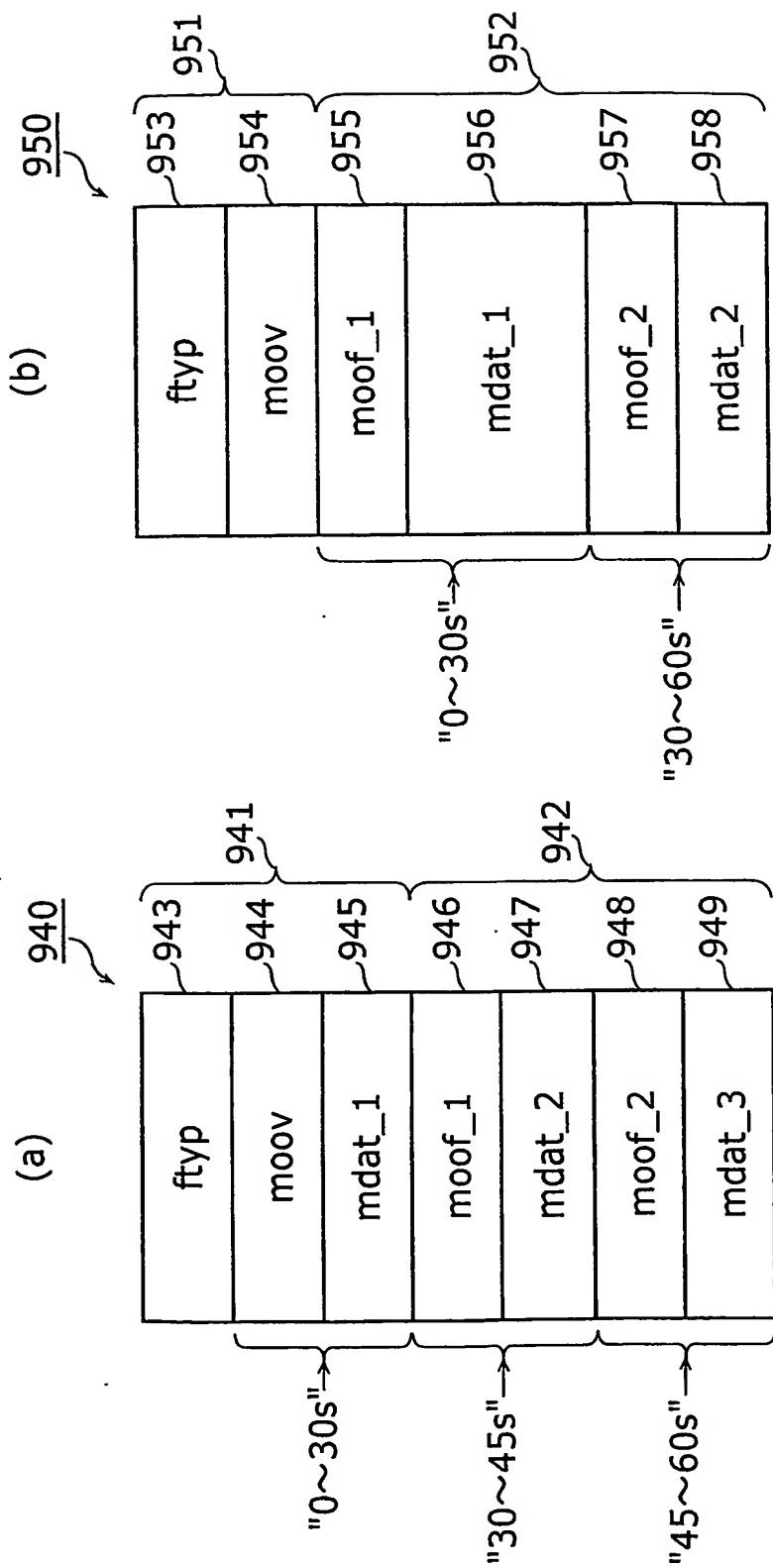


図8

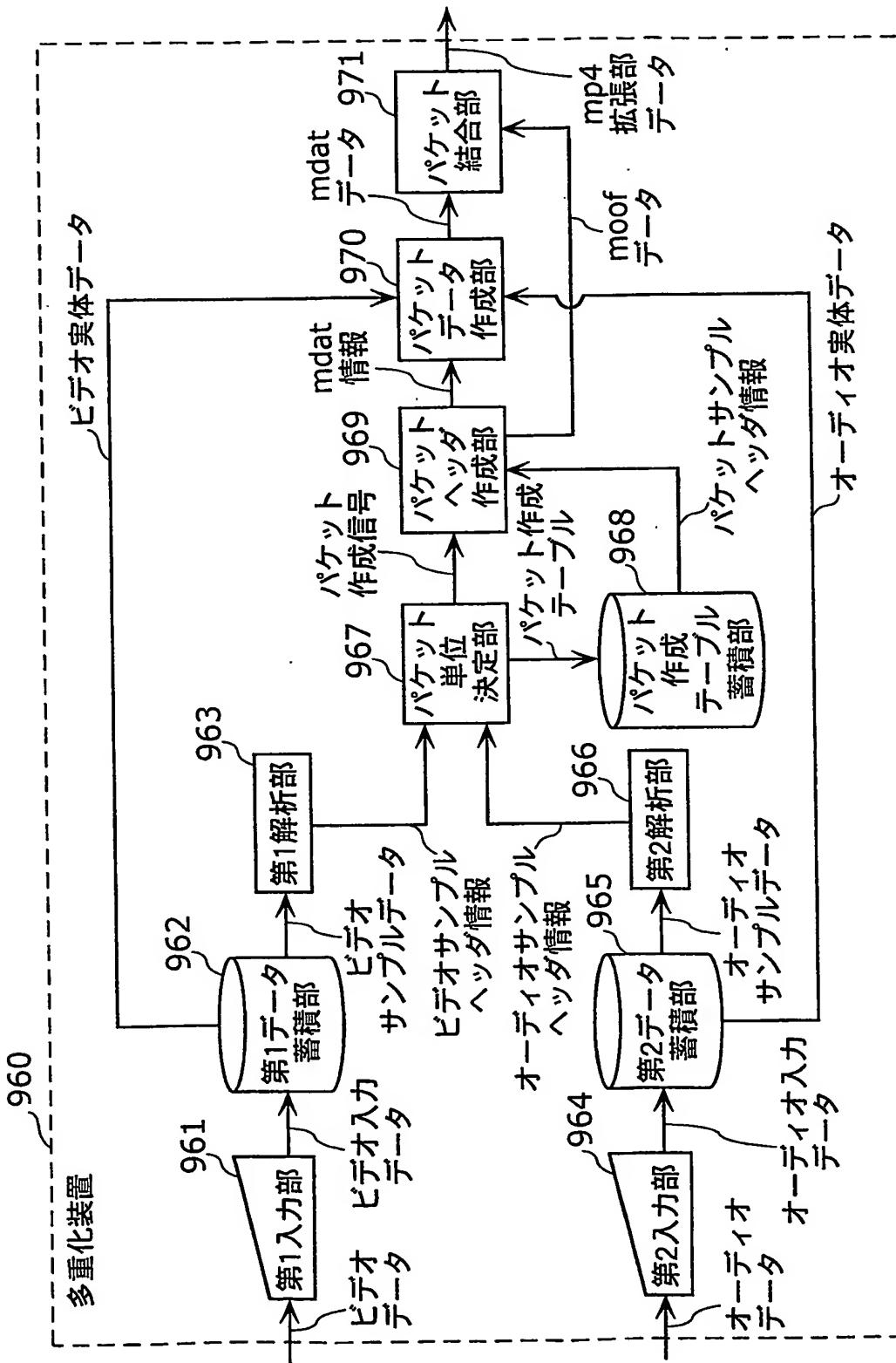


図9

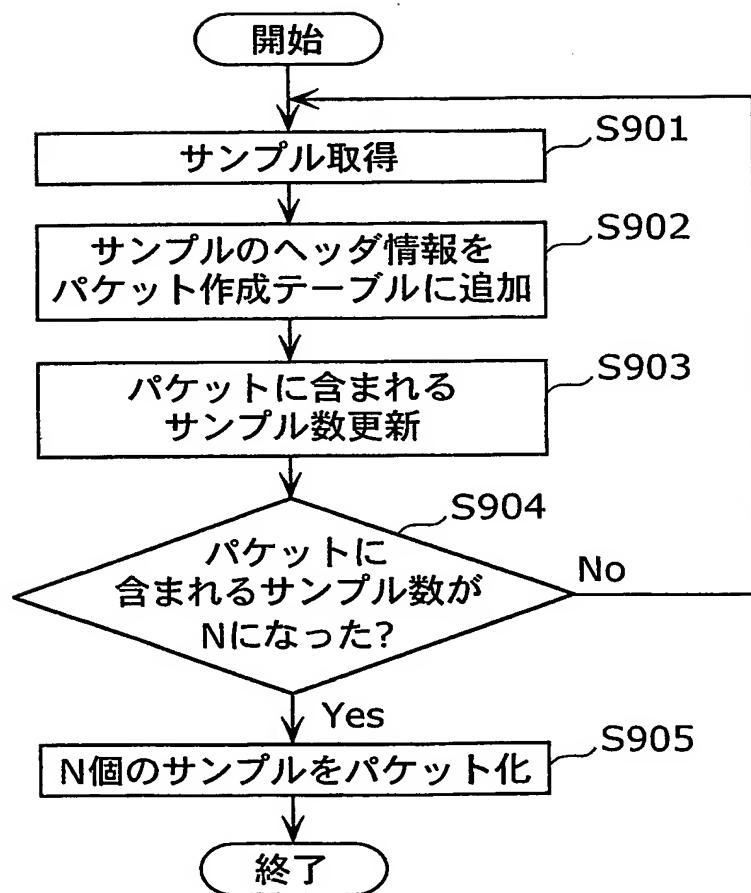


図10

968a

The diagram shows a table with three rows. The first two rows are labeled 'サンプル1' and 'サンプル2' with arrows pointing to them. The third row is labeled 'サンプルN' with an arrow pointing to it. The table has three columns: 'サンプルのサイズ', 'サンプルの再生時間長', and '画面内符号化フレームフラグ'. The data is as follows:

サンプルのサイズ	サンプルの再生時間長	画面内符号化フレームフラグ
300	30	0
300	30	1
⋮	⋮	⋮
300	30	0

図11

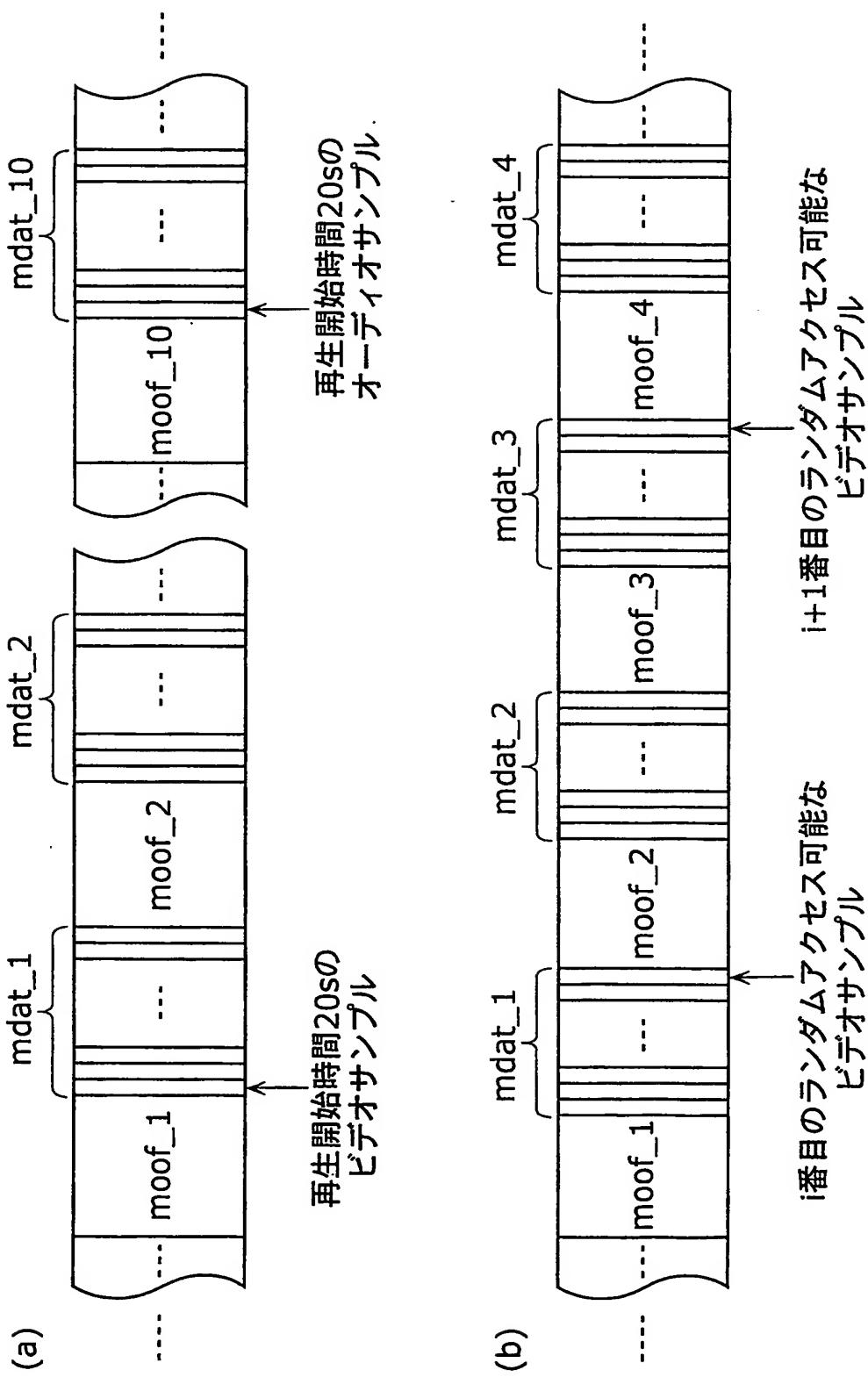


図12

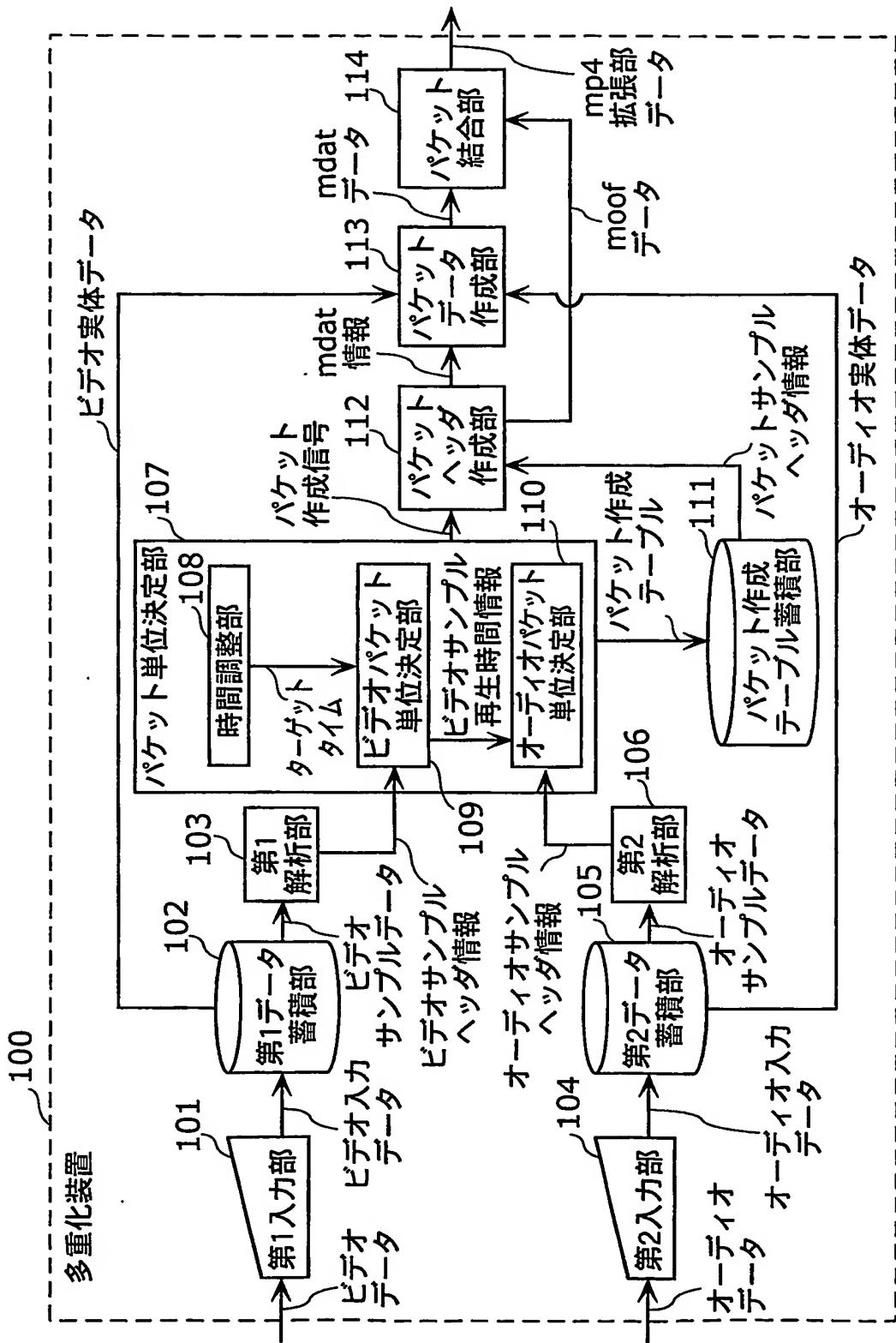


図13

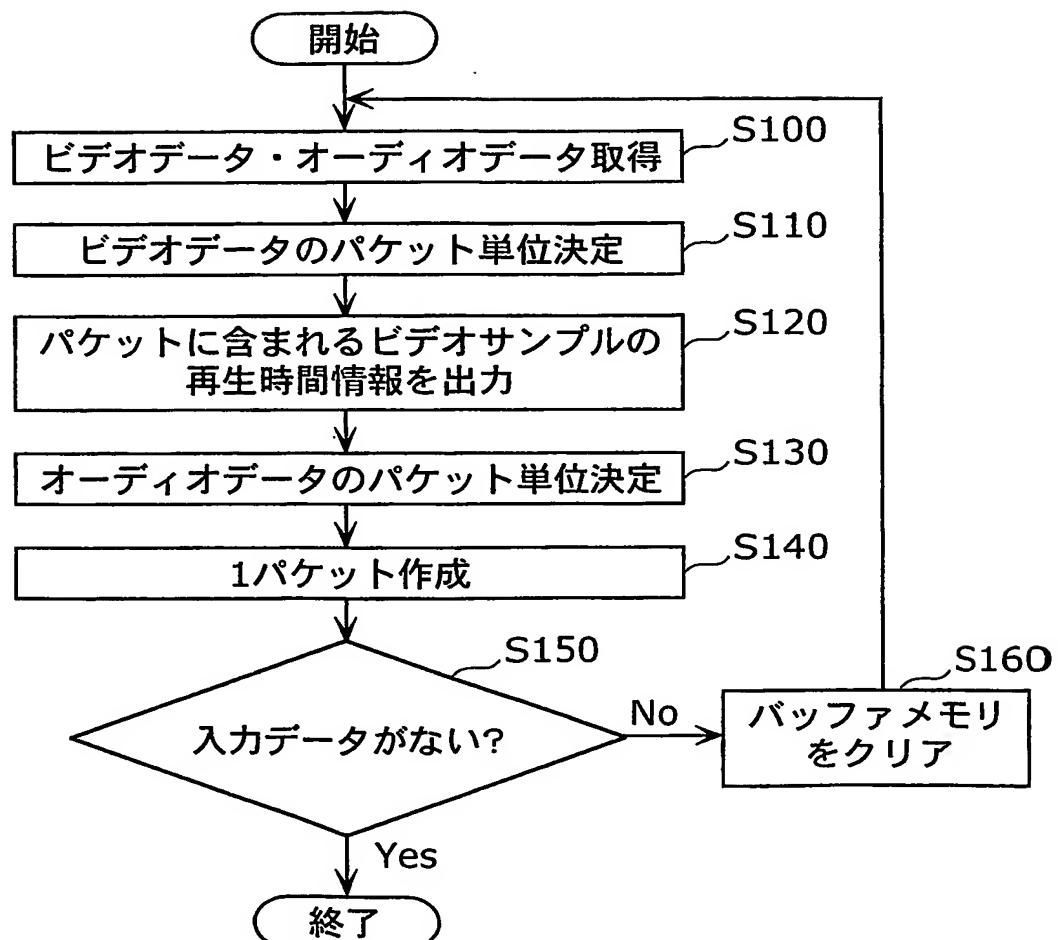


図14

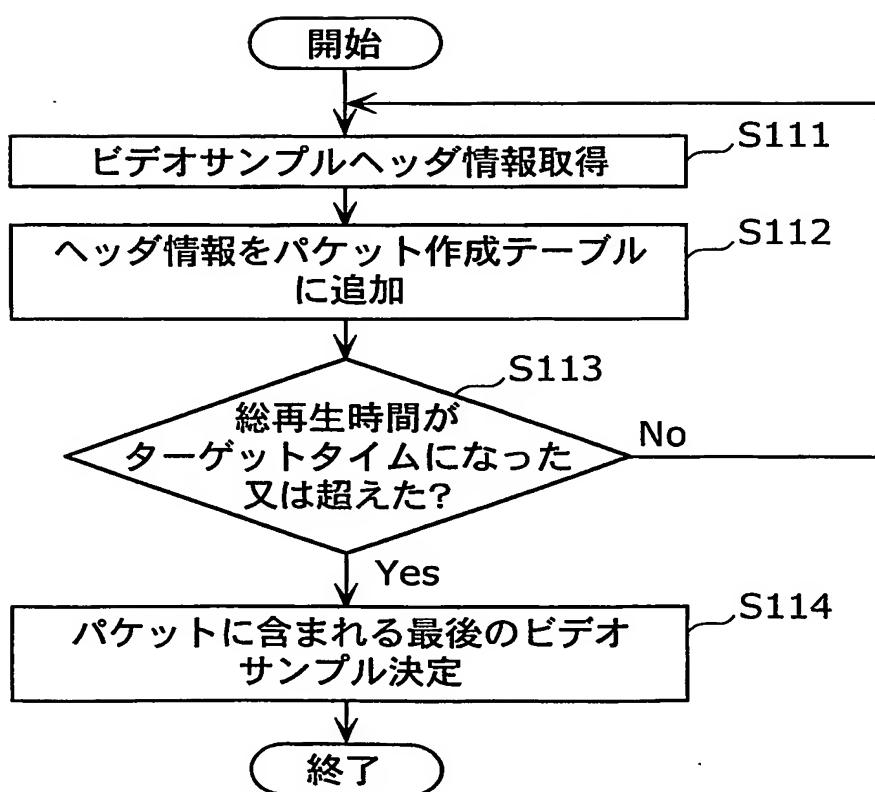


図15

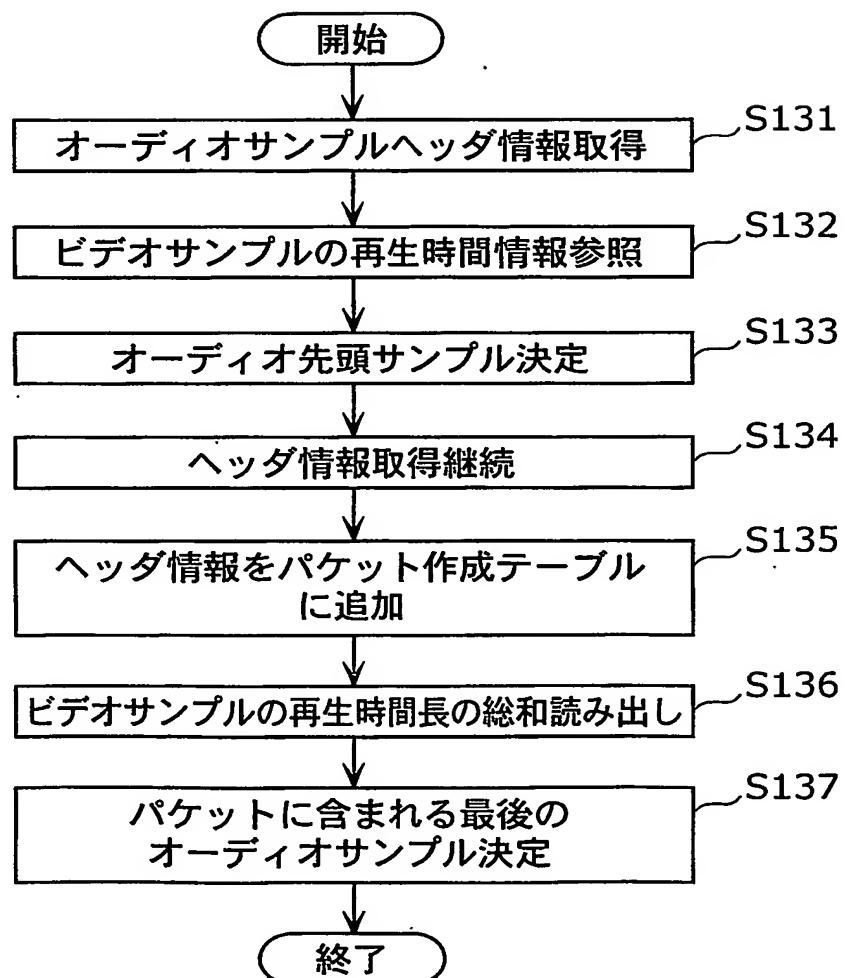


図 16

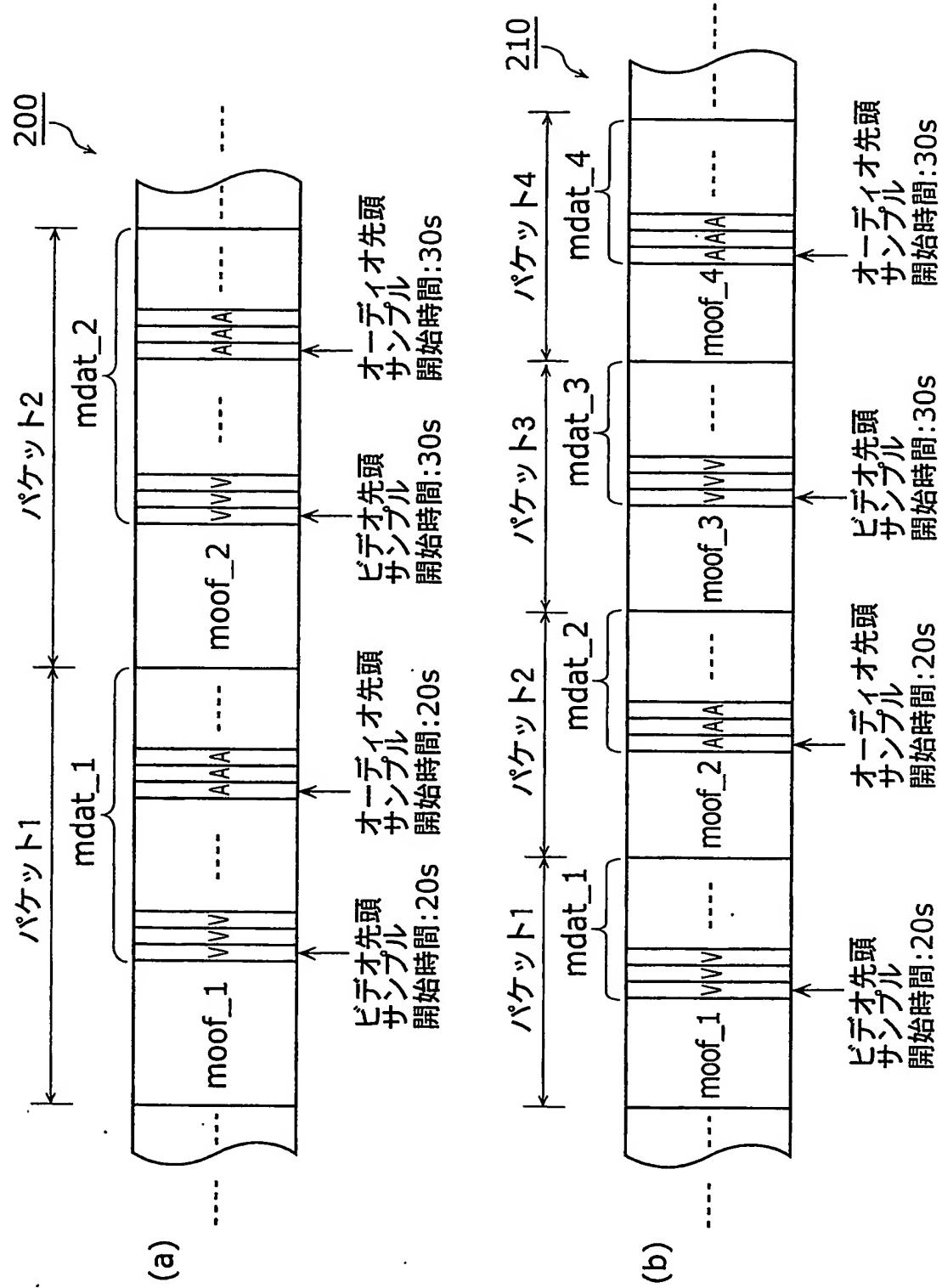


図17

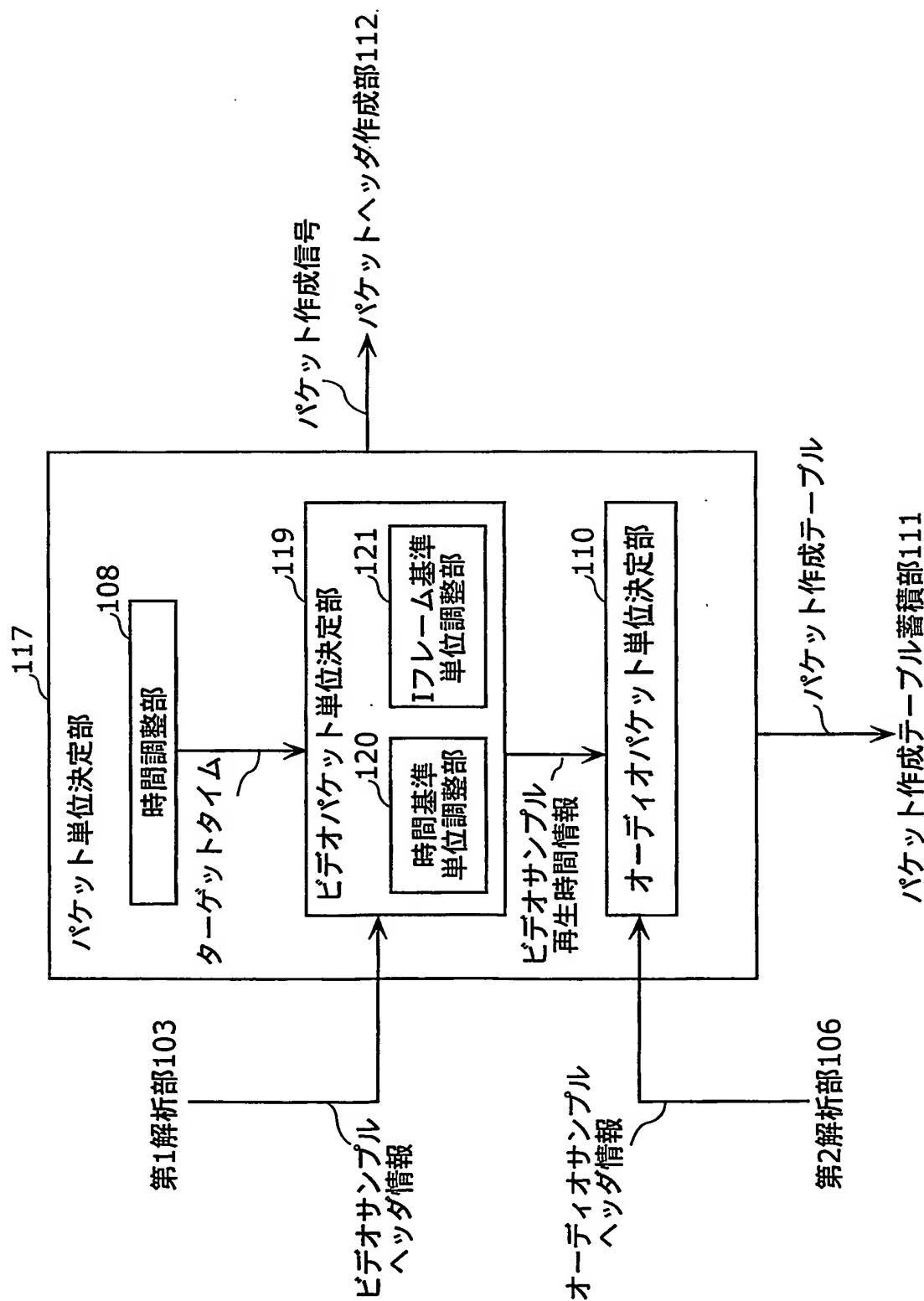


図18

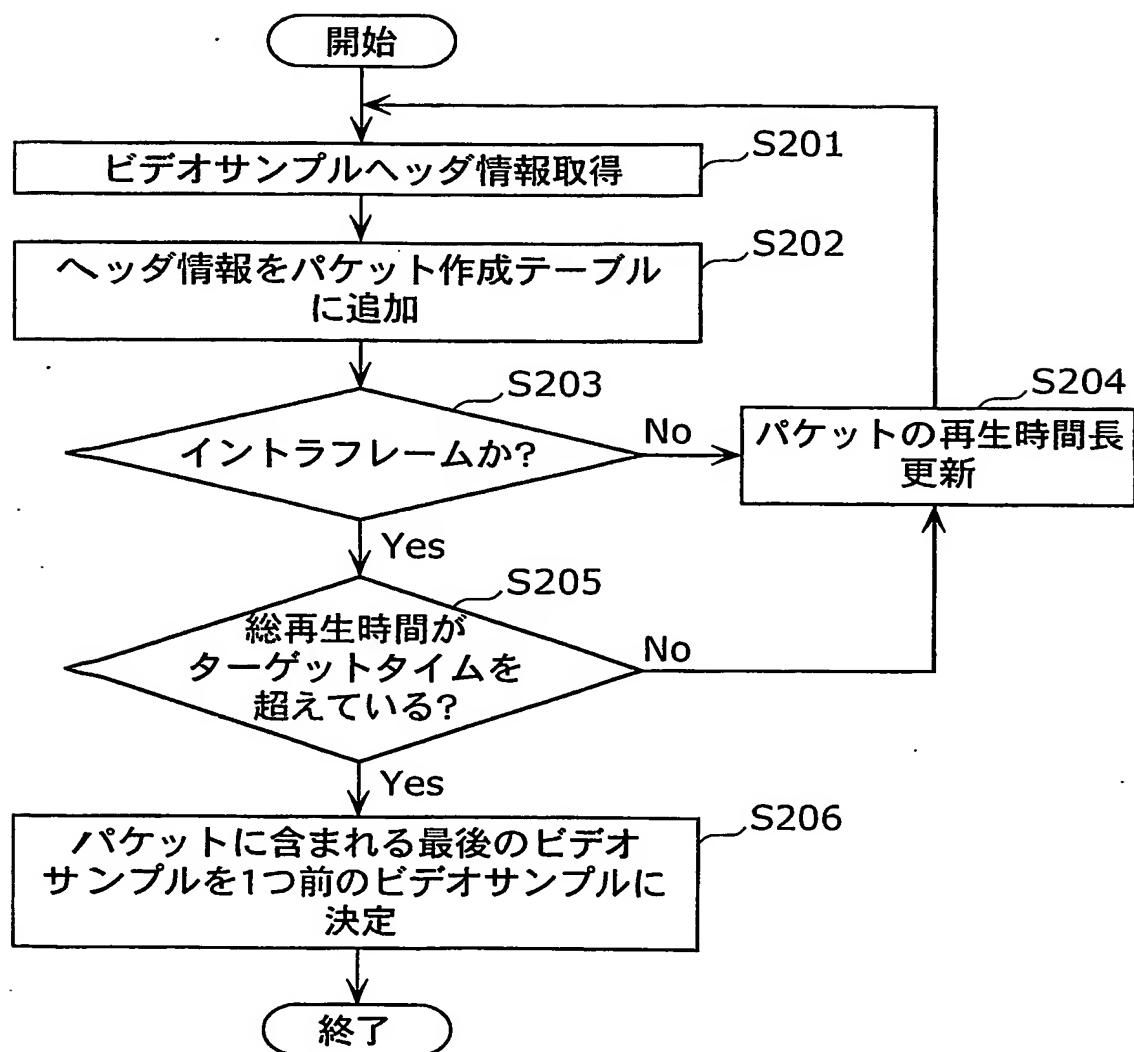


図19

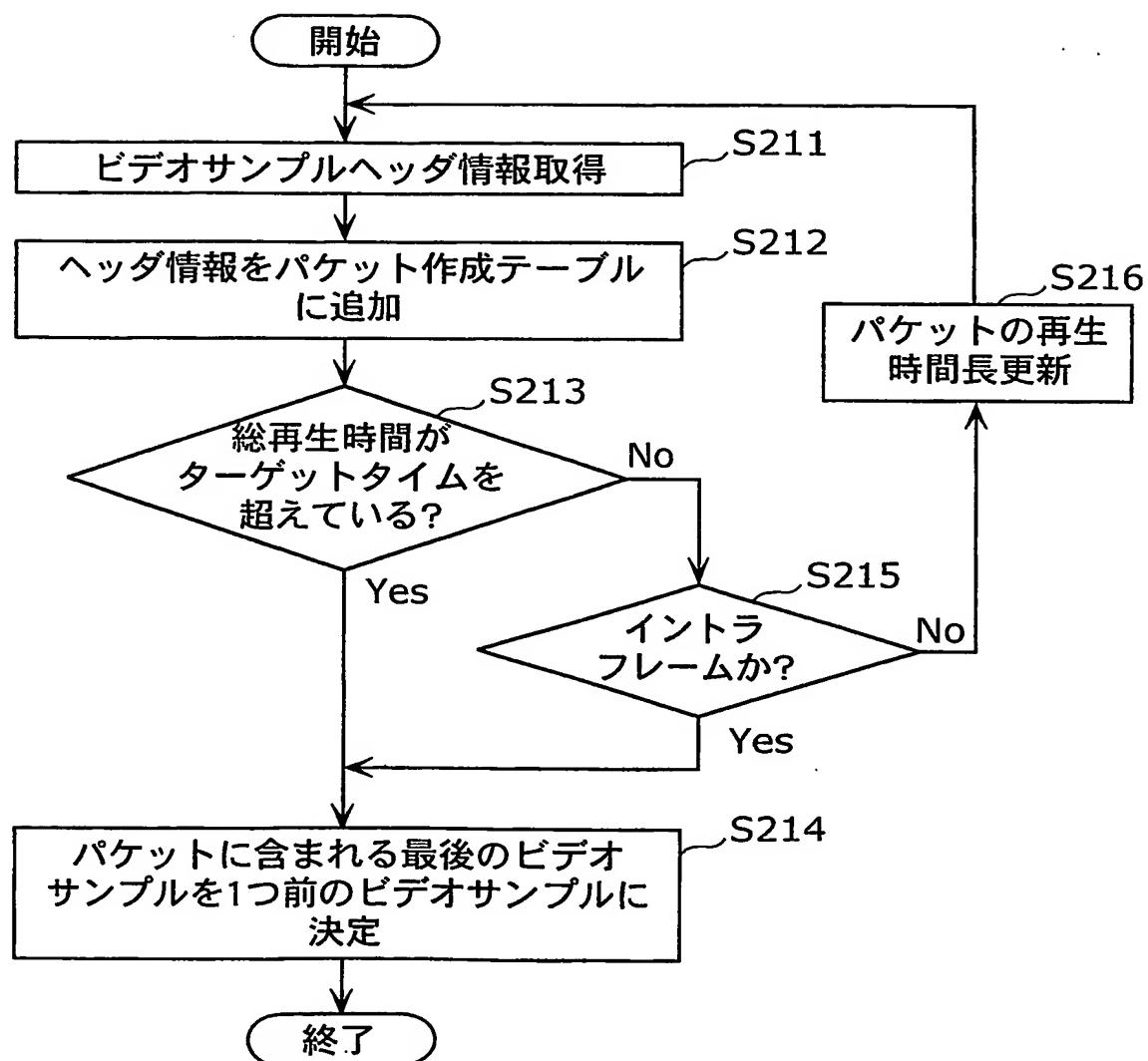


図20

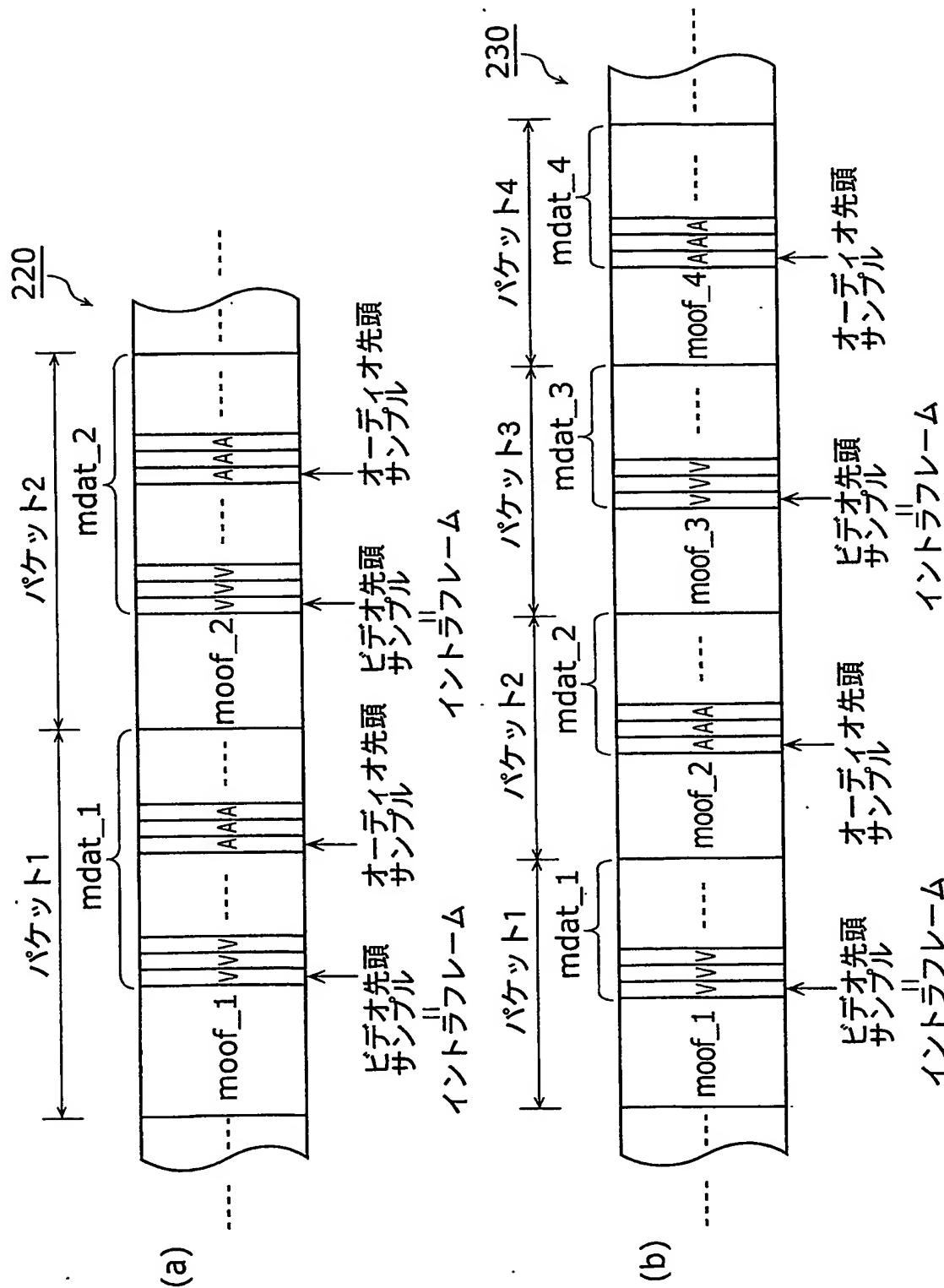


図21

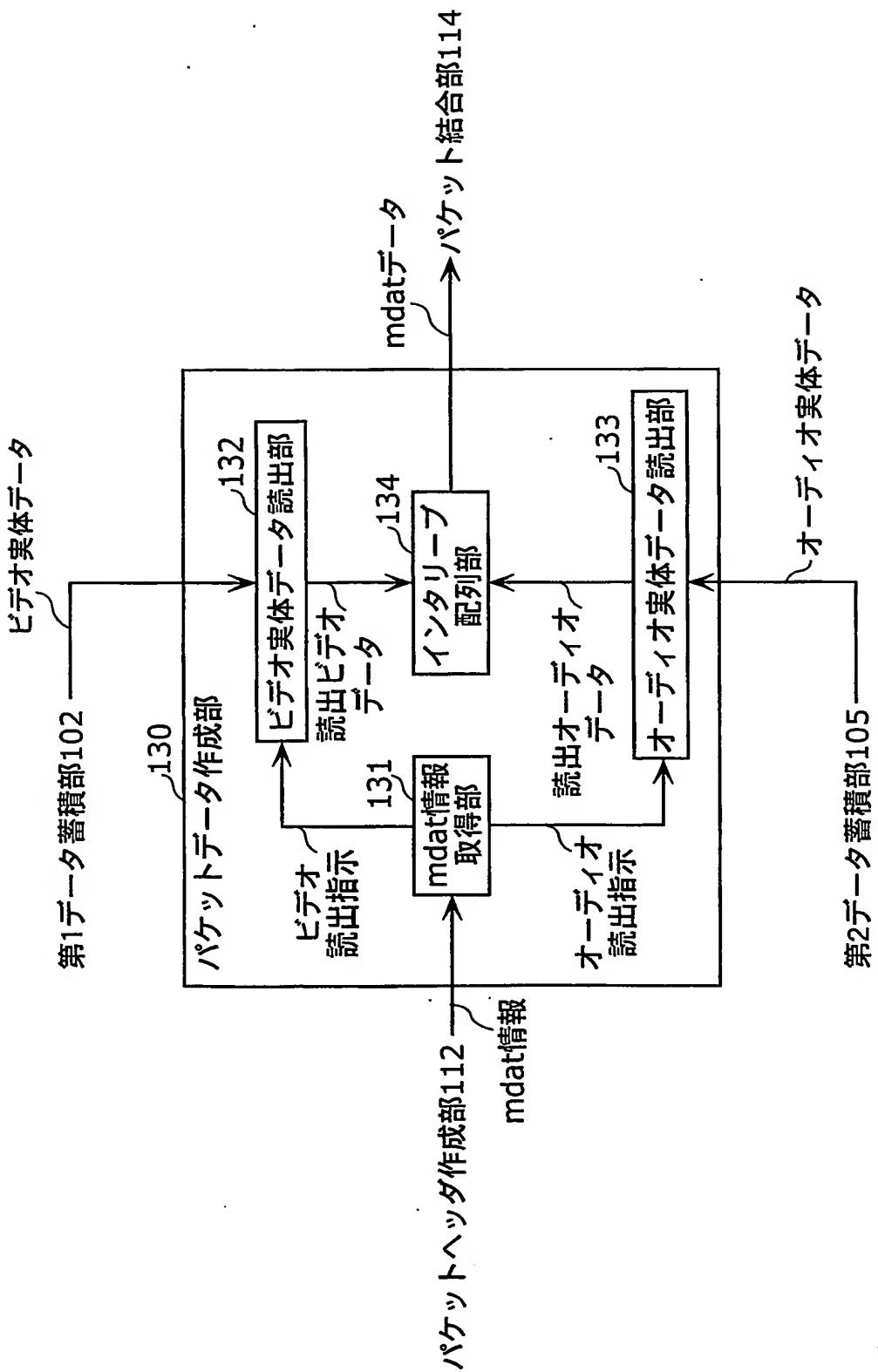


図22

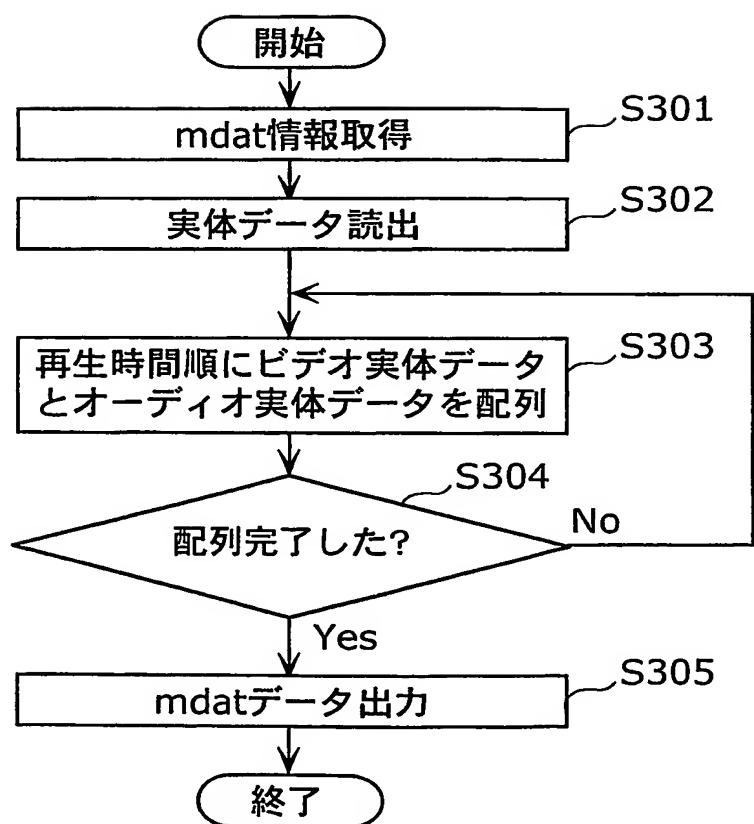


図23

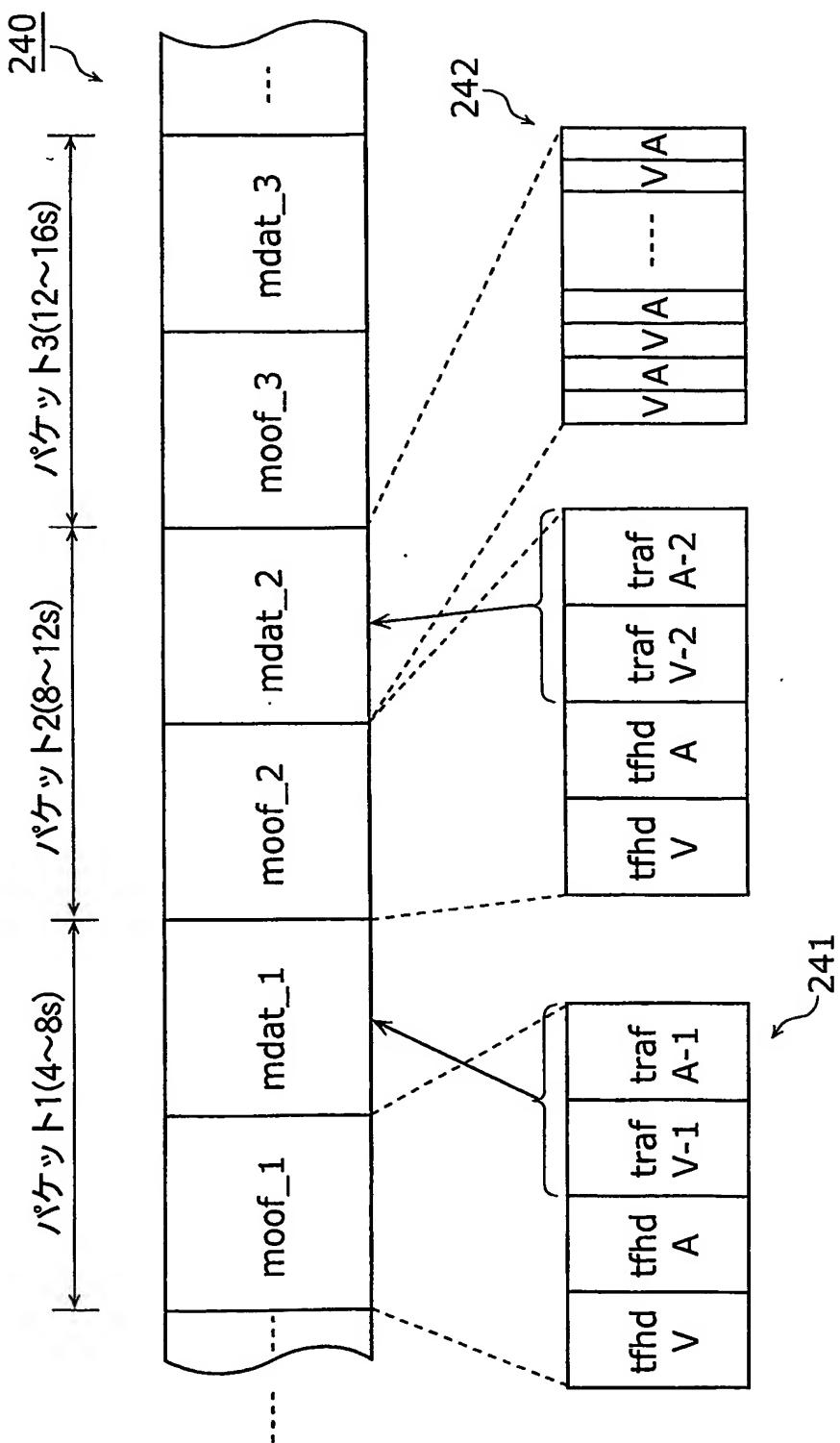


図24

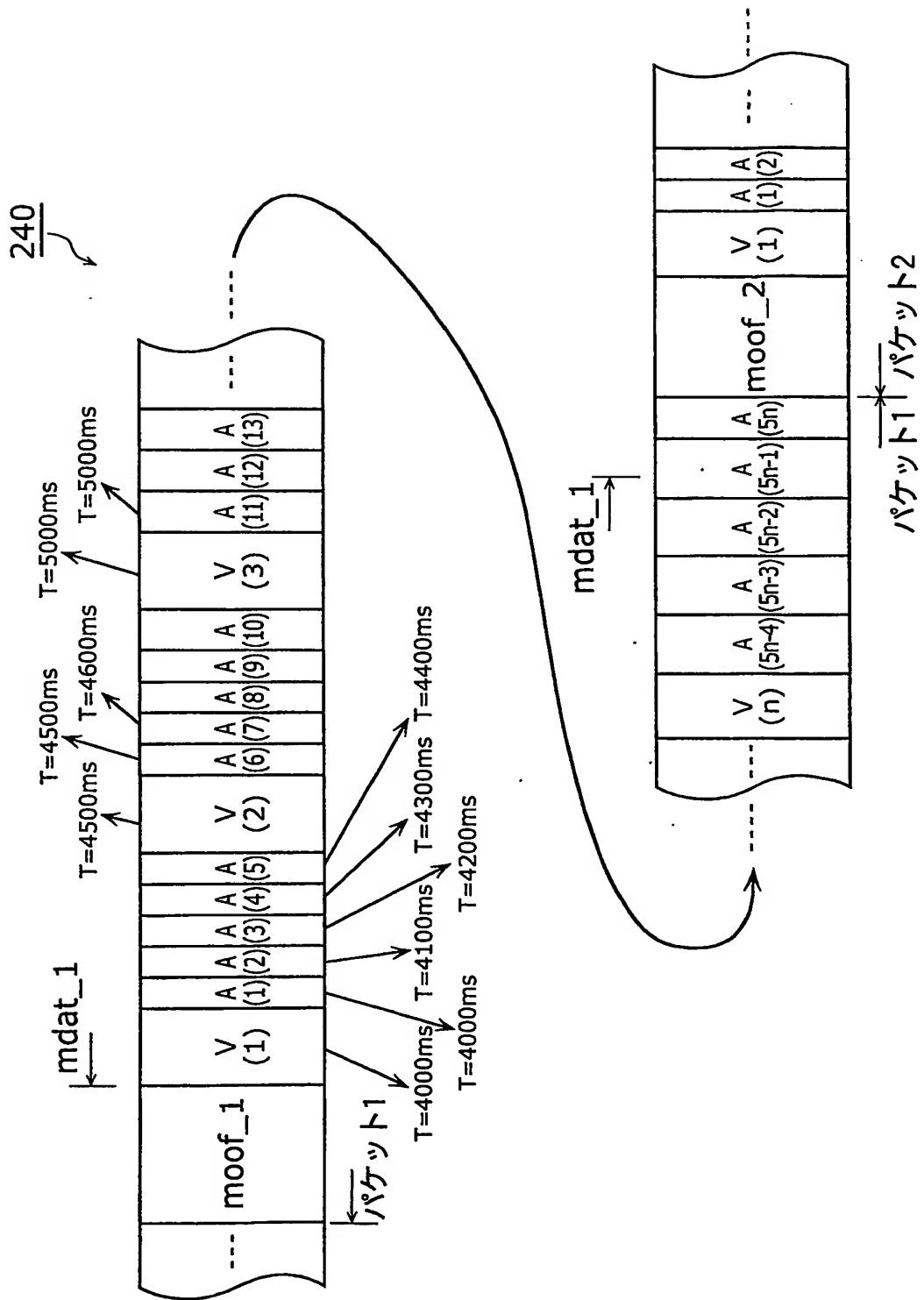


図25

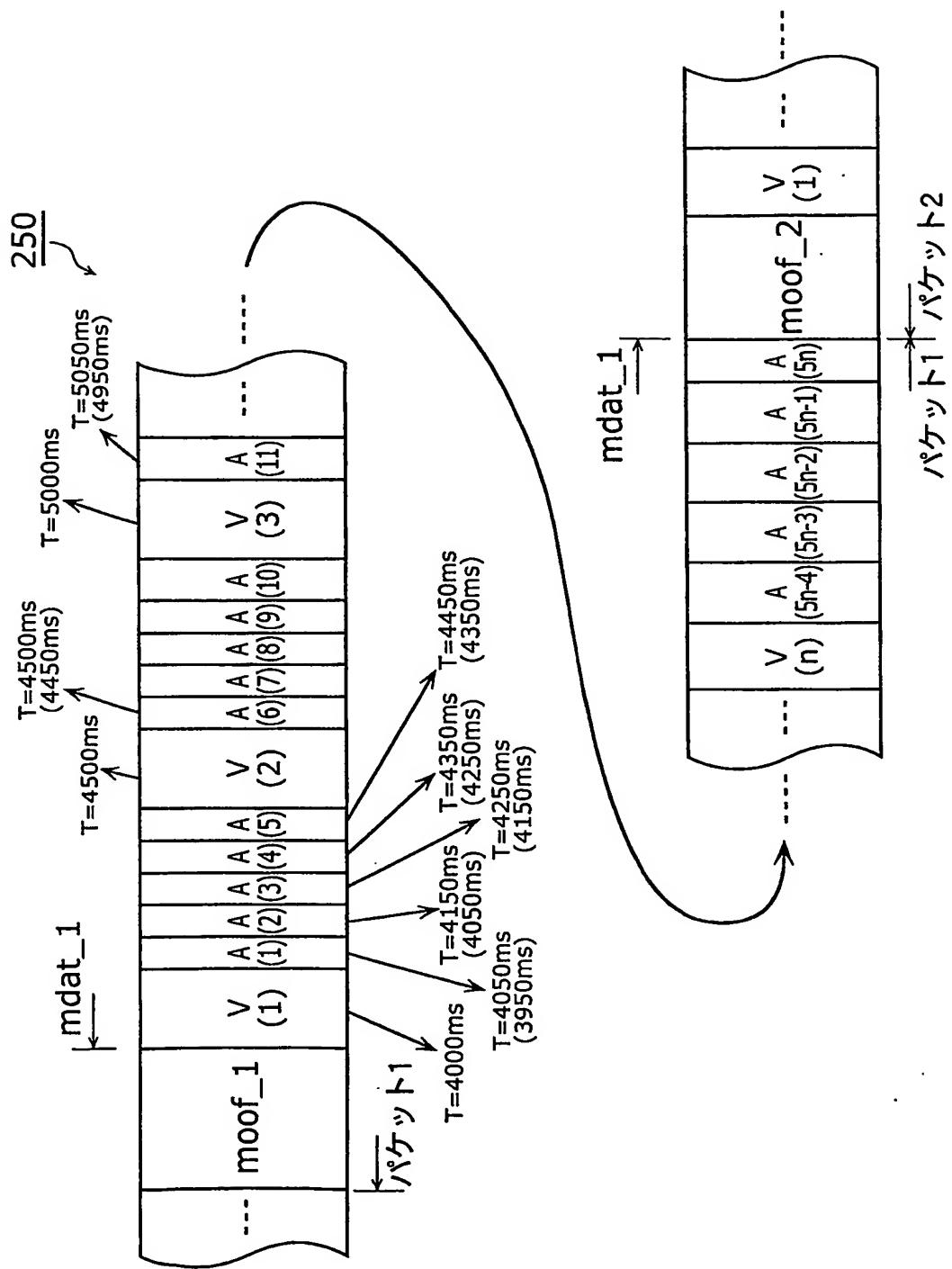


図26

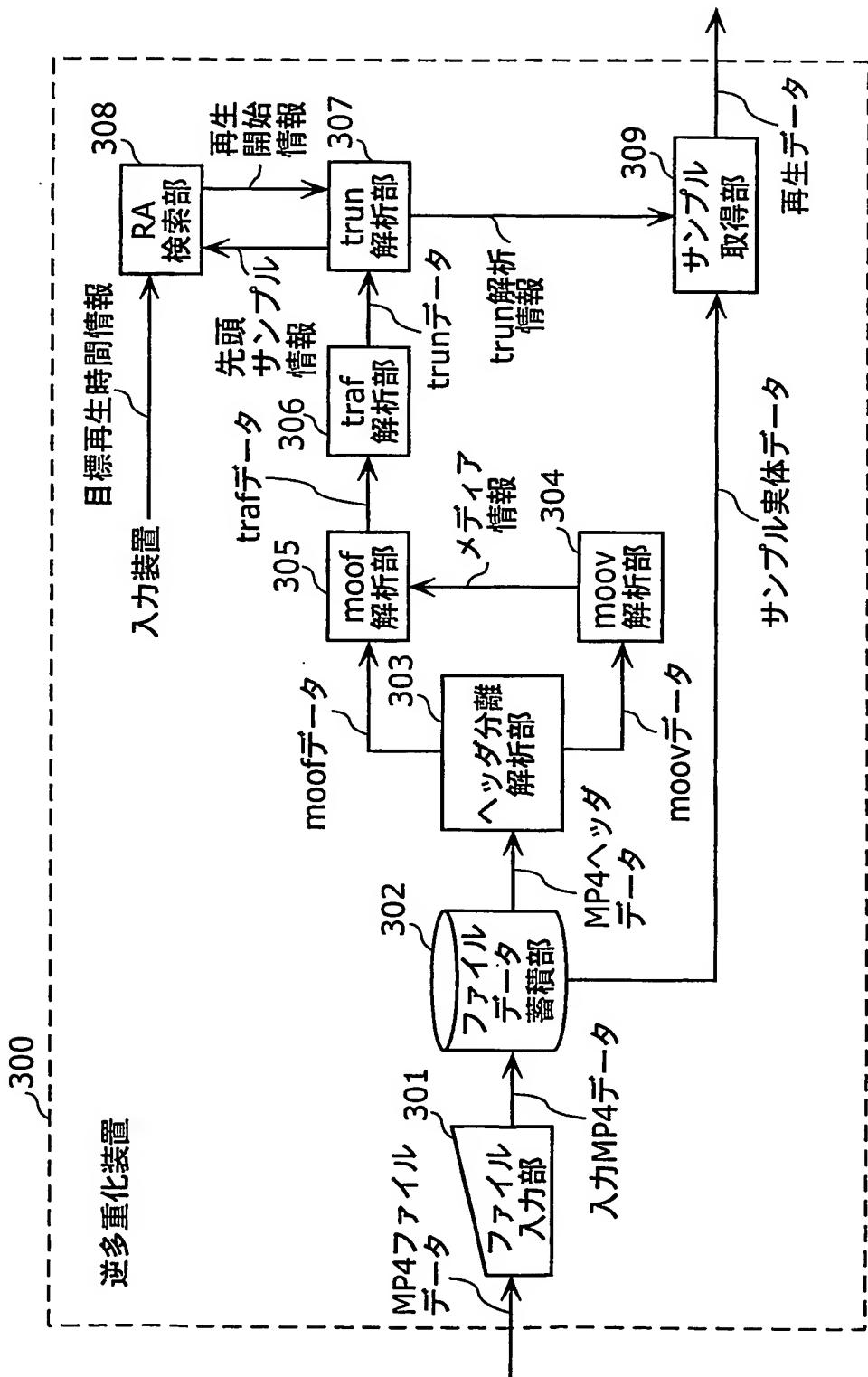


図27

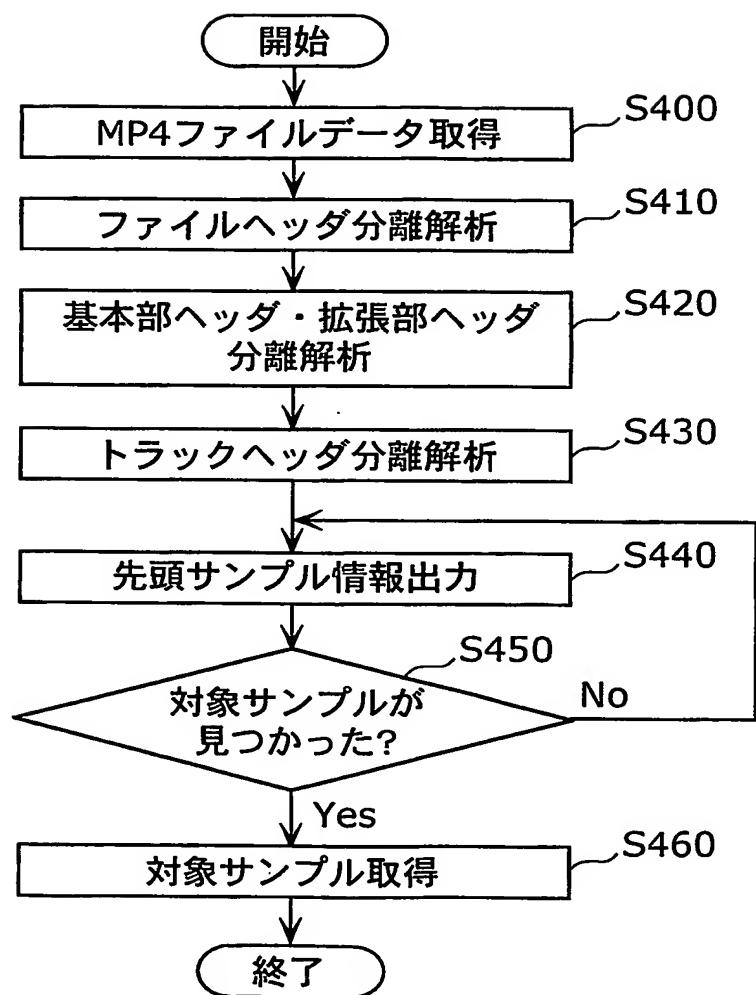
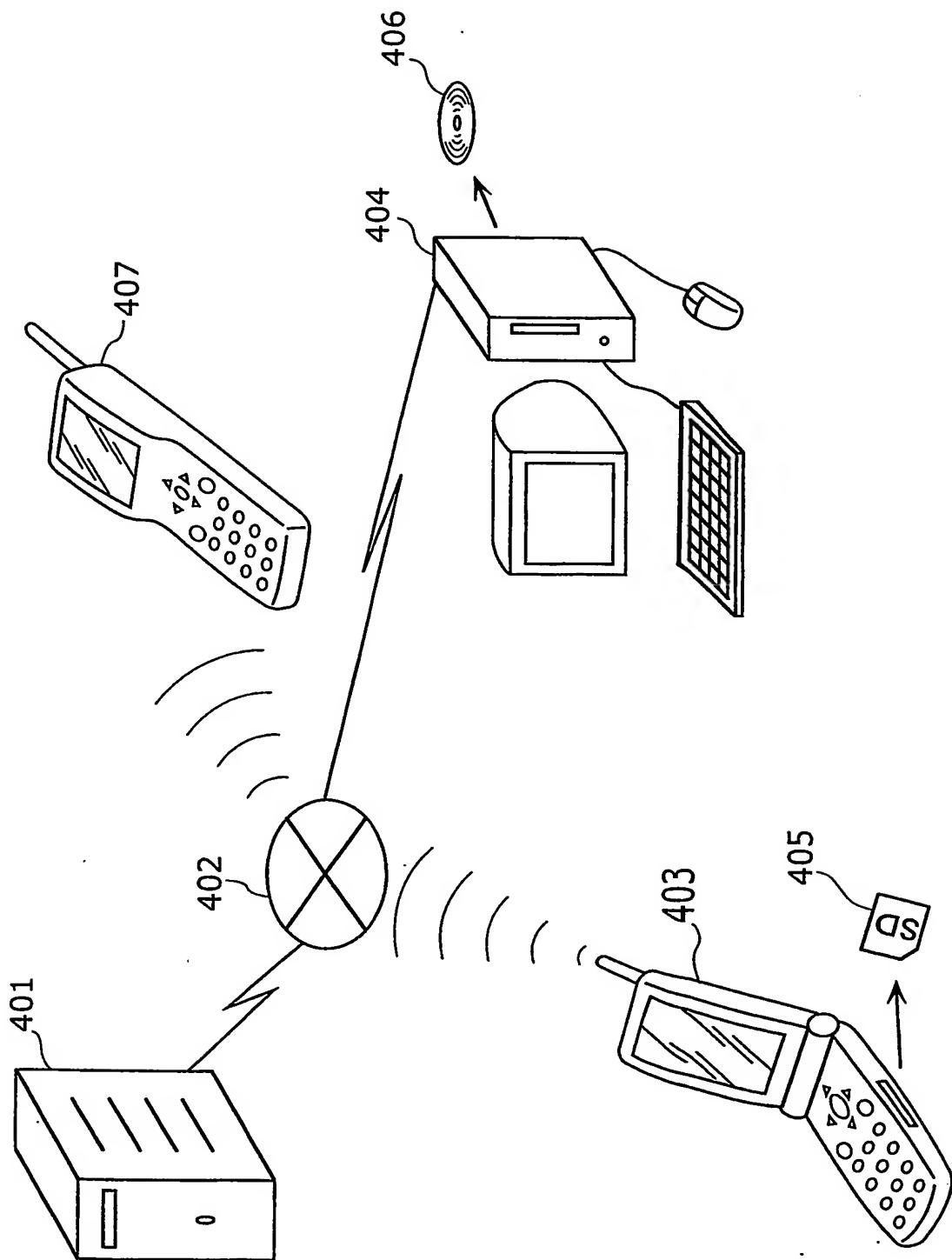


図28



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/07639

A. CLASSIFICATION OF SUBJECT MATTER  
Int.Cl<sup>7</sup> H04N7/08, H04J3/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
Int.Cl<sup>7</sup> H04N7/00-7/088, H04N5/76-5/956, H04N7/12, H04N7/24-7/68Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2003  
Kokai Jitsuyo Shinan Koho 1971-2003 Toroku Jitsuyo Shinan Koho 1994-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 7-38857 A (Pioneer Electronic Corp.), 07 February, 1995 (07.02.95), Par. Nos. [0019] to [0024], [0034] to [0036]; Figs. 5, 8 & EP 634870 B1 & US 557409 A	1-17
X	Hiroshi YASUDA, "Multi Media Fugoka no Kokusai Hyojun", Maruzen Co., Ltd., 30 June, 1991 (30.06. 91), pages 221 to 232	1-4, 8-11, 15-17
Y		5-7, 12-14
X	JP 5-64171 A (Hitachi, Ltd.), 12 March, 1993 (12.03.93), Par. Nos. [0014] to [0020]; Fig. 1 (Family: none)	1-4, 8-11, 15-17
Y		5-7, 12-14

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
---	---

Date of the actual completion of the international search 07 July, 2003 (07.07.03)	Date of mailing of the international search report 29 July, 2003 (29.07.03)
---	--

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP03/07639

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 6-326967 A (Matsushita Electric Industrial Co., Ltd.), 25 November, 1994 (25.11.94), Par. Nos. [0011] to [0015]; Fig. 1 (Family: none)	1-4, 8-11, 15-17 5-7, 12-14
Y	JP 2001-45477 A (NEC Corp.), 16 February, 2001 (16.02.01), Par. Nos. [0008] to [0013]; Fig. 2 (Family: none)	5-7, 12-14

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP03/07639

**Box I Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

The technical feature common to claims 1-17 is that reproduction start time of each sample contained in the medium data is aligned when packetizing the medium data.

However, the aforementioned common feature is not novel since it is disclosed in document JP 7-38857 A (Pioneer Electronic Corporation) and in document YASUDA Hiroshi, "Multimedia hugouka no kokusai hyouzyun (International standard of multimedia coding)", Maruzen, 1991.06.30, p.221-232.

(Continued to extra sheet)

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest  The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP03/07639

Continuation of Box No. II of continuation of first sheet(1)

As a result, the common technical feature makes no contribution over the prior art and cannot be a special technical feature within the meaning of PCT Rule 13.2, second sentence.

Accordingly, it is obvious that claims (1-4, 10, 11, 16, 17), claims (5-7, 12-14), and claims (8, 9, 15) do not satisfy the requirement of unity of invention.

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl' H04N 7/08  
H04J 3/00

## B. 調査を行った分野

## 調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl' H04N 7/00 - 7/088  
H04N 7/12  
H04N 7/24 - 7/68

H04N 5/76 - 5/956

## 最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年  
日本国公開実用新案公報 1971-2003年  
日本国実用新案登録公報 1996-2003年  
日本国登録実用新案公報 1994-2003年

## 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	J P 7-38857 A (パイオニア株式会社) 1995. 02. 07, 段落【0019】～【0024】、【00 34】～【0036】、第5、8図 & E P 634870 B1 & U S 5537409 A	1-17
X	安田 浩, マルチメディア符号化の国際標準, 丸善, 1991. 0 6. 30, p. 221-232	1-4, 8-11, 15-17 5-7, 12-14
Y		

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示す  
もの  
「E」国際出願日前の出願または特許であるが、国際出願日  
以後に公表されたもの  
「L」優先権主張に疑義を提起する文献又は他の文献の発行  
日若しくは他の特別な理由を確立するために引用する  
文献 (理由を付す)  
「O」口頭による開示、使用、展示等に言及する文献  
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

## の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって  
出願と矛盾するものではなく、発明の原理又は理論  
の理解のために引用するもの  
「X」特に関連のある文献であって、当該文献のみで発明  
の新規性又は進歩性がないと考えられるもの  
「Y」特に関連のある文献であって、当該文献と他の1以  
上の文献との、当業者にとって自明である組合せに  
よって進歩性がないと考えられるもの  
「&」同一パテントファミリー文献

国際調査を完了した日 07.07.03	国際調査報告の発送日 29.07.03
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号 100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 菅原 道晴 5 P 3049 電話番号 03-3581-1101 内線 3581

## C (続き) 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP 5-64171 A (株式会社日立製作所) 1993. 03. 12, 段落【0014】～【0020】，第1図 (ファミリーなし)	1-4, 8-11, 15-17
Y		5-7, 12-14
X	JP 6-326967 A (松下電器産業株式会社) 1994. 11. 25, 段落【0011】～【0015】，第1図 (ファミリーなし)	1-4, 8-11, 15-17
Y		5-7, 12-14
Y	JP 2001-45477 A (日本電気株式会社) 2001. 02. 16, 段落【0008】～【0013】，第2図 (ファミリーなし)	5-7, 12-14

## 第I欄 請求の範囲の一部の調査ができないときの意見 (第1ページの2の続き)

法第8条第3項 (PCT17条(2)(a)) の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1.  請求の範囲 \_\_\_\_\_ は、この国際調査機関が調査をすることを要しない対象に係るものである。  
つまり、

2.  請求の範囲 \_\_\_\_\_ は、有意義な国際調査をすることができ程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、

3.  請求の範囲 \_\_\_\_\_ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

## 第II欄 発明の単一性が欠如しているときの意見 (第1ページの3の続き)

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

請求の範囲1-17に共通の事項は、メディアデータに含まれる各サンプルの再生開始時間を揃えて前記メディアデータをパケット化することである。

しかしながら、上記共通の事項は文献JP 7-38857 A (パイオニア株式会社)や文献 安田 浩, マルチメディア符号化の国際標準, 丸善, 1991.06.30, p. 221-232 等に開示されているように周知であるから、新規であるとは認められない。結果として、上記共通の事項は先行技術の域を出ないから、PCT規則13.2の第2文の意味において、上記共通の事項は特別な技術的特徴ではない。

よって、請求の範囲 (1-4, 10, 11, 16, 17), (5-7, 12-14), (8, 9, 15) は発明の単一性の要件を満たしていないことが明らかである。

- 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
- 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
- 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったので、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
- 出願人が必要な追加調査手数料を期間内に納付しなかったので、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

## 追加調査手数料の異議の申立てに関する注意

- 追加調査手数料の納付と共に出願人から異議申立てがあった。
- 追加調査手数料の納付と共に出願人から異議申立てがなかった。